SUPPLEMENTO

AL VOLUME XII, SERIE X

DEL

NUOVO CIMENTO

A CURA DELLA SOCIETÀ ITALIANA DI FISICA

1959

2° Trimestre

PRINTED IN ITALY

NICOLA ZANICHELLI EDITORE
BOLOGNA

Digitized by the Internet Archive
in 2024

https://archive.org/details/supplemento-al-nuovo-cimento_1959_12_1

SUPPLEMENTO
AL VOLUME XII, SERIE X, DEL
NUOVO CIMENTO
A CURA DELLA SOCIETÀ ITALIANA DI FISICA

1959 2° Trimestre N. 1

On the Algebra of Angular Momentum.

A. MECKLER

Division of Physical Sciences, Department of Defense - Baltimore, College Park, Md.

(ricevuto il 16 Luglio 1958)

CONTENTS. — **1.** Introduction. – **2.** Tensor notation and a fundamental theorem. – **3.** Products of components. – **4.** Properties of the multipole tensors. – **5.** Spherical tensors. – **6.** Structure of the finite algebra. – **7.** Structure of the sub-algebra. – **8.** The super algebra. – **9.** Calculus of an intrinsic angular momentum. – **10.** An illustrative problem. – **11.** The Tchebichef polynomials and Clebsch-Gordan coefficients. – APPENDIX I: Reduction of a symmetric tensor. – APPENDIX II: The linear independence of multipole tensor components. – APPENDIX III: The number of independent multipole tensor components. – APPENDIX IV: Proof that $D_{ij\ldots}^{(\mathcal{S}+1)}[\boldsymbol{g}]=0$. – APPENDIX V: The Tchebichef polynomials. – APPENDIX VI: Evaluation of $a_{ij\ldots n}^{(n)}b_{ij\ldots n}^{(n)}$.

1. – Introduction.

The inquiry of this paper begins with the commutation rule for angular momentum and proceeds from it to analyse the algebra generated by the components of an angular momentum vector. (An algebra, over a field, is a system of elements and the operations of addition, multiplication, and scalar multiplication. Addition is commutative and associative. Multiplication is distributive with respect to addition and, for the angular momentum algebra, associative ([1,2])). It is remarkable that the commutation rule suggests and permits a demarcation of the elements of the algebra into linearly independent

([1]) L. E. DICKSON: *Algebras and Their Arithmetics* (New York, 1938).
([2]) D. E. LITTLEWOOD: *A University Algebra* (London, 1950), chap. XIV.

sets of tensor operators. These tensors are related to the forms introduced by RACAH and others ([3,4]). The algebraic presentation directs attention to aspects of these tensors which reveal the structure of the angular momentum algebra, and it suggests manipulations which inevitably lead to an operator calculus which effects the immediate computation of observable quantities, transition probabilities rather than probability amplitudes.

2. – Tensor notation and a fundamental theorem.

The manipulations to come are best handled by tensorial methods. The space is three dimensional and the summation convention is adopted: each dummy index appearing twice is to be summed over the values 1, 2, and 3. A very valuable entity makes it first appearance in the expression of the commutation rule. This is the completely antisymmetric unit pseudotensor of the third rank ([5,6]), written as ε_{ijk}. A component of ε_{ijk} changes sign under the interchange of any two indices, and the non-vanishing components are phased by setting ε_{123} equal unity. Two rules obeyed by ε_{ijk} are easily proved:

$$\varepsilon_{ij\sigma}\varepsilon_{\sigma kp} = \delta_{ik}\delta_{jp} - \delta_{ip}\delta_{jk} , \tag{1}$$

$$\varepsilon_{i\mu\nu}\varepsilon_{\mu\nu i} = 2\delta_{ij} , \tag{2}$$

where, as usual,

$$\delta_{ij} = 0 , \qquad i \neq j,$$

$$= 1 , \qquad i = j.$$

The commutation rule can be written as

$$S_i S_j - S_j S_i = (S_i, S_j) = i\varepsilon_{ij\sigma} S_\sigma . \tag{3}$$

Multiplication (and summation) by ε_{kij} yields an alternative form

$$\varepsilon_{i\mu\nu}(S_\mu, S_\nu) = 2iS_i , \tag{4}$$

or

$$\varepsilon_{i\mu\nu} S_\mu S_\nu = iS_i . \tag{5}$$

Eq. (3) is similarly derivable from Eq. (5).

([3]) A. R. EDMONDS: *Angular Momentum in Quantum Mechanics* (Princeton, 1957).
([4]) M. E. ROSE: *Elementary Theory of Angular Momentum* (New York, 1957).
([5]) L. LANDAU and E. LIFSHITZ: *The Classical Theory of Fields* (Cambridge, Mass., 1951).
([6]) J. L. SYNGE and A. SCHILD: *Tensor Calculus* (Toronto, 1952).

In the construction of a linear algebra, it is necessary to find a set of basal units, a set of elements in which any element of the algebra can be expanded. The components of $\boldsymbol{S}$ are a good beginning, for they are linearly independent. It is a fact that if

$$a_\lambda S_\lambda = b \,,$$

where a_λ and b are members of the field (*), then

$$a_\lambda = 0 \,,$$

$$b = 0 \,.$$

For, on the expression

$$a_\lambda S_\lambda = b$$

perform a commutation with S_ν, getting

$$\varepsilon_{\mu\lambda\sigma} a_\lambda S_\sigma = 0 \,.$$

Perform another commutation, with S_ν:

$$\text{(6)} \qquad \begin{cases} \varepsilon_{\mu\lambda\sigma}\varepsilon_{\sigma\nu k} a_\lambda S_k = 0 \,, \\ a_\lambda S_\lambda \delta_{\mu\nu} - a_\nu S_\mu = 0 \,. \end{cases}$$

Contract this last expression:

$$3a_\lambda S_\lambda - a_\lambda S_\lambda = 0 \,,$$

$$a_\lambda S_\lambda = 0 = b \,.$$

Going back to Eq. (6):

$$a_\nu S_\mu = 0 \,.$$

Division by the member of the field is permissable except in the case that

$$a_\nu = 0 \,.$$

The proof indicates more than the linear independence of the angular momentum components. Actually, in the writing of the expression

$$a_\lambda S_\lambda = b$$

(*) For purposes of this paper, the field need be no more than that of the complex numbers.

there is the tacit assumption of the existence of a modulus, a unit element, in the algebra. Henceforth, let the algebra have associated with it a modulus. The theorem just proved concerns the linear independence of the three components of $\boldsymbol{S}$ and the modulus.

3. – Products of components.

The search for basal units of higher powers of angular momentum components can be carried on in this way. The elementary quadratic form in the components of $\boldsymbol{S}$ is S_iS_j. This product is not a « pure » or irreducible quadratic, inasmuch as

$$S_iS_j = \frac{1}{2}(S_iS_j + S_jS_i) + \frac{1}{2}(S_iS_j - S_jS_i) = \frac{1}{2}(S_iS_j + S_jS_i) + \frac{i}{2}\varepsilon_{ij\sigma}S_\sigma .$$

The symmetrized form, $S_iS_j + S_jS_i$, seems to be the thing to consider. In fact, for a product of any degree, the fully symmetrized form, only, need be considered. The factors (effectively the indices) in a product of n components of $\boldsymbol{S}$ can be re-arranged into any of the $n!$ permutations on n symbols by successive commutations. Each commutation gives a re-arranged product of degree n, plus a homogeneous polynominal of one degree less. By the addition of the $n!$ re-arrangements it is possible to write

$$n!\, S_iS_j \ldots S_n = (\text{sum of all permutations of } S_iS_j \ldots S_n) +$$
$$+ (\text{a linear combination of products of lesser degree}) .$$

The second degree product has been expressed as a sum of the symmetric forms of second and first degree. By induction the product of n components of $\boldsymbol{S}$ can be expressed as a sum of the symmetric tensors of degree n and lower. As notation I shall use

$$\sum^{(n)}_{ij\ldots n} = \| S_iS_j \ldots S_n , \tag{7}$$

where the symbol $\|$ indicates that the form to the right of the symbol is to be brought to full symmetry in all its non-dummy indices.

From the definition, Eq. (7), it should be clear that

$$\sum^{(n)}_{ij\ldots n} = \| \sum^{(n-1)}_{ij\ldots} S_n = \| S_i \sum^{(n-1)}_{j\ldots n} , \tag{8}$$

with the convention that $\sum^{(0)} = 1$, the modulus. As for the commutator of

$\sum^{(n)}$ with S, the symmetrization does not hinder the application of the rule

$$(A, BC) = (A, B)C + B(A, C)$$

so that

$$(S_\mu, \sum^{(n)}_{ij\ldots n}) = i \underset{ij\ldots}{\|} \varepsilon_{\mu i\sigma} \sum^{(n)}_{\sigma j\ldots n} . \tag{9}$$

The indices to be symmetrized are here indicated beneath the double bar for further clarity. Eq. (9) can be looked upon as a generalization of the basic commutation rule, Eq. (3). The $\sum^{(n)}$, however, do not admit a generalization of the commutation rule expressed as

$$\varepsilon_{i\mu\nu}(S_\mu, S_\nu) = 2iS_i ,$$

for multiplication of both sides of Eq. (9) by $\varepsilon_{\lambda\mu i}$ gives

$$\left\{\begin{aligned} &\varepsilon_{\lambda\mu i}(S_\mu, \sum^{(n)}_{ij\ldots n}) = i\varepsilon_{\lambda\mu i} \underset{ij\ldots}{\|} \varepsilon_{\mu i\sigma} \sum^{(n)}_{\sigma j\ldots n} = i(\varepsilon_{\lambda\mu i}\varepsilon_{\mu i\sigma} \sum^{(n)}_{\sigma j\ldots n} + \varepsilon_{\lambda\mu i} \underset{j\ldots n}{\|} \varepsilon_{\mu j\sigma} \sum^{(n)}_{\sigma ik\ldots n}) = \\ &\qquad\qquad = i(2 \sum^{(n)}_{\lambda j\ldots n} + (n-1) \sum^{(n)}_{\lambda j\ldots n} - \underset{j\ldots n}{\|} \delta_{j\lambda} \sum^{(n)}_{\sigma\sigma\ldots n}) \\ &\varepsilon_{\lambda\mu i}(S_\mu, \sum^{(n)}_{ij\ldots n}) = i[(n+1) \sum^{(n)}_{\lambda j\ldots n} - \underset{j\ldots n}{\|} \delta_{j\lambda} \sum^{(n)}_{\sigma\sigma\ldots n}] . \end{aligned}\right. \tag{10}$$

The contracted tensors disturb the harmony of the commutation rules. In fact, they prevent the linear independence of the $\Sigma^{(n)}$ tensors, as an attempt at a proof will show. There is a compulsion to modify the $\Sigma^{(n)}$ so that a contraction yields a null tensor. To illustrate:

$$\Sigma_{ij} = S_iS_j + S_jS_i ,$$

$$\Sigma_{\mu\mu} = 2S^2 .$$

Let

$$D_{ij} = \Sigma_{ij} - \tfrac{2}{3}S^2\delta_{ij} .$$

Then

$$D_{\mu\mu} = 0 .$$

D_{ij} should be generalized. The task is to preserve the fully symmetric homogeneous polynomial character, yet assure a vanishing contraction. An attempt at a construction is to begin with

$$D^{(n)}_{ij\ldots n} = \underset{}{\|} D^{(n-1)}_{ij\ldots} S_n$$

following Eq. (8) for $\Sigma^{(n)}_{ij\ldots n}$, and then to subtract off $\frac{1}{2}n(n-1)$ diagonal terms:

$$D^{(n)}_{ij\ldots n} = \| D^{(n-1)}_{ij\ldots} S_n - \| \delta_{ij} B^{(n-2)}_{k\ldots n} ,$$

where the $B^{(n)}$ are to be fully symmetric and also of vanishing contraction. Given that the contraction of $D^{(n-1)}$ vanishes, the condition that $D^{(n)}$ behave likewise is

$$0 = 2 D^{(n-1)}_{\mu k\ldots} S_\mu - 3 B^{(n-2)}_{k\ldots n} - 2(n-2) B^{(n-2)}_{k\ldots n} ,$$

which can be solved to give

$$B^{(n-2)}_{k\ldots n} = \frac{2}{2n-1} D^{(n-1)}_{\mu k\ldots n} S_\mu .$$

Therefore, let a multipole tensor (*) be defined by

$$D^{(n)}_{ij\ldots n} = \left\| D^{(n-1)}_{ij\ldots} S_n - \frac{2}{2n-1} \right\| \delta_{ij} D^{(n-1)}_{\mu k\ldots} S_\mu , \tag{11}$$

with

$$D^{(0)} = 1$$

and

$$D^{(1)}_i = S_i .$$

There is a direct way of modifying a symmetric tensor so that the modification has a vanishing contraction. It is a messy business and is demonstrated in Appendix I. The explicit reduction procedure applied to $\Sigma^{(n)}_{ij\ldots n}$, which obeys the iterative formula of Eq. (8), must lead to the $D^{(n)}_{ij\ldots n}$ which obeys the iterative formula of Eq. (11). The explicit procedure shows that $D^{(n)}_{ij\ldots n}$ is a linear combination of $\Sigma^{(n)}_{ij\ldots n}$ and its contractions. A p-times contracted $\Sigma^{(n)}_{ij\ldots n}$ can be man-handled into lesser degree symmetric tensors multiplied by functions of S^2. Therefore, $D^{(n)}_{ij\ldots n}$ is a linear combination of the symmetric tensors of degree n and lower, the coefficients being functions of S^2. The roles can be reversed by and inductive argument based on

$$\Sigma_{ij} = D_{ij} + \tfrac{2}{3} S^2 \delta_{ij} .$$

That is, a symmetric tensor, $\Sigma^{(n)}_{ij\ldots n}$, can be expressed as a linear combination of multipole tensors, but the coefficients are functions of S^2. Any product of angular momentum components can be expressed as such a linear combination of multipole tensors, but the coeffcients are not simply numbers. The multipole tensors are not quite basal units at the moment.

(*) L. Landau and E. Lifshitz ([5]), p. 104.

4. – Properties of the multipole tensors.

$D^{(n)}_{ij...n}$ is a homogeneous polynomial in the components of $\boldsymbol{S}$. It is as true for $D^{(n)}_{ij...n}$ as it is for $\Sigma^{(n)}_{ij...n}$ that the operation of commutation with S_μ is distributive over the indices, each index in turn being replaced by $i\varepsilon_{\mu\lambda}D^{(n)}_{...\lambda...}$. The only point to be checked is the commutation with factors like $S^2\delta_{ij}$. But this is all right, as one can write

$$(S_\mu, S^2\delta_{ij}) = iS^2(\varepsilon_{\mu i\sigma}S_{\sigma j} + \varepsilon_{\mu j\sigma}S_{i\sigma}) = iS^2(\varepsilon_{\mu ij} + \varepsilon_{\mu ji}) ,$$

which is equal to zero, as it should be. Therefore, one form of the commutation rule is (*)

$$(S_\mu, D^{(n)}_{ij...n}) = i \underset{ij...}{\|} \varepsilon_{\mu i\sigma}D^{(n)}_{\sigma j...n} . \tag{12}$$

The multipole tensors were introduced to permit the transformation of Eq. (10) into

$$\varepsilon_{i\mu\nu}(S_\mu, D^{n}_{ij...n}) = i(n+1)D^{(n)}_{ij...n} \tag{13}$$

the second form of the commutation rule.

Take Eq. (12) and bring it to full symmetry in all the indices:

$$\underset{ij...}{\|}(S_\mu, D^{(n)}_{ij...n}) = i \underset{ij...}{\|} \varepsilon_{\mu i\sigma}D^{(n)}_{\sigma j...n} .$$

Symmetrization of the antisymmetric unit tensors will cause cancellation, so that

$$\underset{ij...}{\|}(S_\mu, D^{(n)}_{ij...n}) = 0 . \tag{14}$$

From this equation, coupled with the condition of vanishing contraction, it follows that

$$(S_\mu, D^{(n)}_{\mu j...n}) = 0 . \tag{15}$$

Eqs. (14) and (15) allow flexibility in the generating formula for the multipole tensors, Eq. (11), in that the multiplication can be performed from either side.

(*) This equation could be derived by induction on Eq. (11).

To get an identity of tantalizing appearance, replace $D^{(n)}_{\nu j\ldots n}$ in the left-hand side of Eq. (13) by its formula given by Eq. (13):

$$\varepsilon_{i\mu\nu}\varepsilon_{\nu\lambda\sigma}(S_\mu, (S_\lambda, D^{(n)}_{\sigma j\ldots n})) = -(n+1)^2 D^{(n)}_{ij\ldots n},$$

$$(D^{(n)}_{\mu j\ldots n}, (S_\mu, S_i)) + (S_\mu, (S_\mu, D^{(n)}_{ji\ldots n})) = (n+1)^2 D^{(n)}_{ij\ldots n}.$$

The last line results from the use of the identity

$$(A, (B, C)) + (B, (C, A)) + (C, (A, B)) = 0$$

and the application of Eq. (15). Going on:

$$-i\varepsilon_{i\sigma\mu}(S_\sigma, D^{(n)}_{\mu j\ldots n}) + (S_\mu, (S_\mu, D^{(n)}_{ij\ldots n})) = (n+1)^2 D^{(n)}_{ij\ldots n}$$

and finally

$$(S_\mu, (S_\mu, D^{(n)}_{ij\ldots n}) = n(n+1) D^{(n)}_{ij\ldots n}. \tag{16}$$

Because $D^{(n)}_{ij\ldots n}$ commutes with S^2, the last equation can be written in two other ways:

$$S_\mu D^{(n)}_{ij\ldots n} S_\mu = [S^2 - \tfrac{1}{2}n(n+1)] D^{(n)}_{ij\ldots n} \tag{17}$$

and

$$S_\mu (S_\mu, D^{(n)}_{ij\ldots n}) = \tfrac{1}{2}n(n+1) D^{(n)}_{ij\ldots n}. \tag{18}$$

A clue to the most important information contained in the multipole tensors is in the calculation:

$$S_\mu D_{\mu j} = S_\mu (S_\mu S_j + S_j S_\mu - \tfrac{2}{3} S^2 \delta_{\mu j}) =$$

$$= S^2 S_j + S_\mu S_j S_\mu - \tfrac{2}{3} S^2 S_j = \tfrac{1}{3} S^2 S_j + (S^2 - 1) S_j$$

because of Eq. (17). In final form

$$S_\mu D_{\mu j} = \tfrac{4}{3}(S^2 - \tfrac{3}{4}) D_j.$$

It can be expected that the tensor defined by $S_\mu D^{(n)}_{\mu j\ldots n}$ will be proportional to $D^{(n-1)}_{j\ldots n}$, because that tensor must be a fully symmetric homogeneous polynomial in the components of $\boldsymbol{S}$, and it must have a vanishing contraction. A proof by induction will confirm this expectation and will reveal the signi-

ficant proportionality factor. It turns out that

$$S_\mu D^{(n)}_{\mu j\ldots n} = \frac{n^2}{2n-1}\left[S^2 - \frac{1}{4}(n^2-1)\right] D^{(n-1)}_{j\ldots n} . \tag{19}$$

Set $n = 2\mathcal{S}+1$ to emphasize the bracketed factor as

$$[S^2 - \mathcal{S}(\mathcal{S}+1)] .$$

The recursion relation for the multipole tensors can now be re-written as

$$D^{(n)}_{ij\ldots n} = \left\| S_i D^{(n-1)}_{j\ldots n} - \frac{2(n-1)^2}{(2n-1)(2n-3)}\left[S^2 - \frac{n}{4}(n-2)\right] \right\|_{ij} \delta_{ij} D^{(n-2)}_{k\ldots n} . \tag{20}$$

Assurance should be given that the identities derived here do permit a proof of the linear independence of the multipole tensor components. A lengthy proof appears in Appendix II. The statement is that if

$$\sum_{(n)} a^{(n)}_{ij\ldots n} D^{(n)}_{ij\ldots n} = 0 ,$$

where the scalar tensor $a^{(n)}_{i\ldots n}$ is fully symmetric in its indices and has vanishing contraction, then

$$a^{(n)}_{i\ldots n} = 0 .$$

Of course, any tensor undergoing inner multiplication with a multipole tensor can be brought to full symmetry and modified so as to have vanishing contraction, without adding anything to the product. This statement of the linear independence of the multipole tensor components means that *all* the linear dependence is recognized when account is taken of the symmetry of $D^{(n)}_{ij\ldots n}$ and of its vanishing contraction. It is therefore possible to count the number of independent elements of a multipole tensor, and this is done in Appendix III. The number of independent elements of $D^{(n)}_{ij\ldots n}$ is $2n+1$.

5. – Spherical tensors.

The multipole tensors which were introduced here, in the search for an algebra, with no thought of the eigenvalue-eigenvector problem of angular momentum, assume much more significance when they are related to certain tensors which may be typified as spherical representations of the cartesian forms developed here (3), (4). These spherical tensors are generalizations of

the well-known shift operators of angular momentum. These shift operators can be defined in the following invariant way:

Let a_μ be a real, three dimensional vector normalized so that

$$a_\mu a_\mu = 1 \,. \tag{21}$$

Let

$$Y_1^0 = a_\mu S_\mu \,.$$

Define

$$Y_1^1 = a_\mu^+ S_\mu$$

and insist that

$$(Y_1^0,\, Y_1^1) = Y_1^1 \,. \tag{22}$$

The expansion of the last equation is

$$i a_\nu a_\mu^+ \varepsilon_{\nu\mu\sigma} S_\sigma = a_\sigma^+ S_\sigma \,.$$

Coefficients of S_σ can be equated. Therefore

$$\varepsilon_{\sigma\mu\nu} a_\mu^+ a_\nu = i a_\sigma^+ \,. \tag{23}$$

Introduce the antisymmetrical tensor dual to a_μ:

$$d_{\sigma\mu} = \varepsilon_{\sigma\mu\nu} a_\nu = - d_{\mu\sigma} \,.$$

The converse relation

$$a_\mu = \tfrac{1}{2}\, \varepsilon_{\mu\nu\sigma} d_{\nu\sigma}$$

is also true.

Eq. (23) can be re-written as

$$d_{\sigma\mu} a_\mu^+ = i a_\sigma^+ \,. \tag{24}$$

This equation can be interpreted, properly, as an eigenvalue-eigenvector equation, for the normalization condition, Eq. (21), implies that

$$1 = \tfrac{1}{4} \varepsilon_{\mu\nu k} d_{jk} \varepsilon_{\mu\lambda\sigma} d_{\lambda\sigma} = \tfrac{1}{4} (d_{\lambda\tau} d_{\lambda\sigma} - d_{\sigma\lambda} d_{\lambda\sigma}) = - \tfrac{1}{2} \operatorname{Tr} \boldsymbol{d}^2 \,.$$

That is

$$\operatorname{Tr} \boldsymbol{d}^2 = - 2 \,.$$

One eigenvalue of a real, antisymmetrical tensor of the third order must be zero; the other two must be conjugate imaginaries. The condition on the trace of $\boldsymbol{d}^2$ means that the eigenvalues of $\boldsymbol{d}$ are 0, $+i$, and $-i$. By Eq. (24), a^+ is the eigenvector belonging to the eigenvalue $+i$. The eigenvector belonging to the eigenvalue 0 is a_σ, for

$$d_{\mu\sigma}a_\sigma = \varepsilon_{\mu\sigma\lambda}a_\lambda a_\sigma = 0$$

because of the antisymmetrical nature of $\varepsilon_{\mu\sigma\lambda}$.

If a_μ^- denotes the eigenvector belonging to the eigenvalues $-i$, then

$$Y_1^{-1} = a_\mu^- S_\mu$$

satisfies the commutation rule

$$(Y_1^0, Y_1^{-1}) = -Y_1^{-1}. \tag{25}$$

Clearly

$$a_\mu^- \propto (a_\mu)^*$$

and it is also easily proved that

$$a_\mu^+ a_\mu = a_\mu^- a_\mu = a_\mu^+ a_\mu^+ = a_\mu a_\mu^- = 0.$$

Normalization can be arranged so that

$$a_\mu^+ a_\mu^- = 1. \tag{26}$$

With this normalization it can be proved that

$$\varepsilon_{\sigma\mu\nu}a_\mu^+ a_\nu^- = -ia_\sigma \tag{27}$$

so that

$$(Y_1^1, Y_1^{-1}) = Y_1^0. \tag{28}$$

Now $a_\mu^+ a_\nu^- + a_\mu^- a_\nu^+ + a_\mu a_\nu$ is a tensor which acts as the unit tensor on the three linearly independent vectors a_μ, a_μ^+, and a_μ^-. These three vectors span the space, and so

$$a_\mu^+ a_\nu^- + a_\mu^- a_\nu^+ + a_\mu a_\nu = \delta_{\mu\nu}. \tag{29}$$

(This is a statement of the unitary character of the transformation.) By

multiplication

$$S_\mu = a_\mu \boldsymbol{a}\cdot\boldsymbol{S} + a_\mu^+ \boldsymbol{a}^-\cdot\boldsymbol{S} + a_\mu^- \boldsymbol{a}^+\cdot\boldsymbol{S} \tag{30}$$

and

$$S^2 = (\boldsymbol{a}\cdot\boldsymbol{S})^2 + \boldsymbol{a}^+\cdot\boldsymbol{S}\,\boldsymbol{a}^-\cdot\boldsymbol{S} + \boldsymbol{a}^-\cdot\boldsymbol{S}\,\boldsymbol{a}^+\cdot\boldsymbol{S}\,. \tag{31}$$

Eqs. (28) and (31) can be combined to give

$$\boldsymbol{a}^+\cdot\boldsymbol{S}\,\boldsymbol{a}^-\cdot\boldsymbol{S} = \tfrac{1}{2}[S^2 - \boldsymbol{a}\cdot\boldsymbol{S}(\boldsymbol{a}\cdot\boldsymbol{S} - 1)]\,, \tag{32}$$

$$\boldsymbol{a}^-\cdot\boldsymbol{S}\,\boldsymbol{a}^+\cdot\boldsymbol{S} = \tfrac{1}{2}[S^2 - \boldsymbol{a}\cdot\boldsymbol{S}(\boldsymbol{a}\cdot\boldsymbol{S} + 1)]\,. \tag{33}$$

The spherical tensors are generated in this way:
Define

$$Y_N^N = (Y_1^1)^N$$

$$Y_N^{N-1} = (Y_1^1, Y_N^N)$$

$$Y_N^{N-2} = (Y_1^1, Y_N^{N-1})$$

and so on down to Y_N^{-N}. (The normalization is terrible, but that's not important here.) A general commutation rule, easily proved, is

$$(Y_1^0, Y_N^M) = M Y_N^M\,. \tag{34}$$

Eq. (34) leads to the interpretation of Y_N^M as a shift operator which preserves S^2 but shifts the eigenvalue of $\boldsymbol{a}\cdot\boldsymbol{S}$ by M units. An angular momentum system whose highest possible multiplicity is $2\mathcal{S}+1$ can be characterized by the statement

$$Y_{\mathcal{S}+1}^{2\mathcal{S}+1} = 0\,. \tag{35}$$

Now,

$$Y_N^N = a_i^+ \ldots a_N^+ S_i \ldots S_N \propto a_i^+ \ldots a_N^+ D^{(N)}_{ij\ldots N}$$

because $a_i^+ \ldots a_N^+$ is a symmetrical tensor of vanishing contraction. If

$$Y_N^N = 0$$

then

$$a_i^+ \ldots a_N^+ D^{(N)}_{ij\ldots N} = 0\,.$$

The linear independence of the multipole tensor components would demand that

$$a_i \dots a_N = 0$$

but this is absurd. $Y_N^N = 0$ must mean that $D^{(N)}_{ij\dots N} = 0$. The converse is obviously true.

6. – Structure of the finite algebra.

What are the consequences of constraining the angular momentum algebra? Let it be that

$$D^{(N)}_{ij\dots N} = 0 \tag{36}$$

but that no lower tensor vanishes. Multiply Eq. (36) by S_i:

$$S_i D^{(N)}_{ij\dots N} = 0$$

which by Eq. (19), becomes

$$[S^2 - \tfrac{1}{4}(N^2 - 1)] D^{(N-1)}_{j\dots N} = 0 .$$

$D^{(N-1)}_{j\dots N} = 0$, so it is permissable to perform another multiplication (by S_j) and so on, arriving at

$$\prod_{k=0}^{N-2} \left[S^2 - \frac{(N-k)^2 - 1}{4} \right] S_i = 0 , \tag{37}$$

and, by one more multiplication, at

$$\prod_{k=0}^{N-1} \left[S^2 - \frac{(N-k)^2 - 1}{4} \right] = 0 . \tag{38}$$

Eq. (38) can be called the Hamilton-Cayley equation of S^2. It is well known what to do with such an equation [8]. Let

$$\alpha(k) = \frac{(N-k)^2 - 1}{4} , \tag{39}$$

and let

$$E(k) = \prod_{\substack{\rho=0 \\ \neq k}}^{N-1} \frac{S^2 - \alpha(\rho)}{\alpha(k) - \alpha(\rho)} . \tag{40}$$

By Eq. (38)

$$S^2E(k) = E(k)S^2 = \alpha(k)E(k)\,. \tag{41}$$

$E(k)$ is a polynomial in S^2 of degree $(N-1)$. $\sum_k E(k)$ is a polynomial in S^2, defined over the field, of degree $(N-1)$. At the N points, $(k=0, \ldots, N-1)$, $\sum_k E(k)$ has the value 1. This is a statement concerning the field coefficients in the polynomial; it has nothing to do with the nature of angular momentum. Therefore,

$$\sum_k E(k) = 1\,. \tag{42}$$

The summation of Eq. (41) yields

$$S^2 = \sum \alpha(k)E(k)\,. \tag{43}$$

It follows from Eq. (40) and (41) that

$$E(k)E(k') = 0\,, \qquad k \neq k', \tag{44}$$

and that

$$E^2(k) = E(k)\,. \tag{45}$$

Eqs. (44) and (45) characterize the E's as orthogonal idempotents.

Any component of $\boldsymbol{S}$ can be resolved in this fashion, for

$$S_i = \sum E(k)S_i$$

or

$$S_i = \sum h_i(k)\,, \tag{46}$$

where

$$h_i(k) = E(k)S_i\,. \tag{47}$$

The interesting thing is that $h_i(k)$ is itself an angular momentum:

$$(h_i, h_j) = (ES_i, ES_j) = E(S_i, S_j) = i\varepsilon_{ij\sigma}ES_\sigma = i\varepsilon_{ij\sigma}h_\sigma\,.$$

The above steps are possible because E commutes with S_i.

From Eq. (47)

$$\left\{\begin{array}{l} h^2 = ES_iES_i = ES^2 \\ h^2(k) = \varkappa(k)E(k). \end{array}\right. \tag{48}$$

Moreover,

$$h_i(k)h_j(k') = 0\,, \qquad k \neq k'. \tag{49}$$

$\boldsymbol{h}$ has a very simple Hamilton-Cayley equation:

$$h^2 h_i = \alpha h_i\,. \tag{50}$$

The $\boldsymbol{h}$ is an angular momentum. Its components can generate a subalgebra in the algebra of $\boldsymbol{S}$. Let a multipole tensor formed with the components of $\boldsymbol{h}$ be denoted as

$$D^{(n)}_{ij\ldots}[\boldsymbol{h}(k)]\,.$$

Then

$$D^{(n)}_{ij\ldots}[\boldsymbol{h}(k)] = E(k)D^{(n)}_{ij\ldots}[\boldsymbol{S}] = 0$$

for $k \geqslant N - n$, because $E(k)$ would then contain at least all the factors necessary to satisfy

$$\prod_{p=0}^{N-n-1}\left[S^2 - \frac{(N-p)^2 - 1}{4}\right] D^{(n)}[\boldsymbol{S}] = 0\,.$$

Therefore,

$$D^{(N-k)}_{ij\ldots}[\boldsymbol{h}(k)] = 0\,.$$

To summarize, the important equations of this section can be written as

$$D^{(2\mathcal{S}+1)}_{ij\ldots}[\boldsymbol{S}] = 0\,, \tag{51}$$

$$S^2 = \sum h(h+1)E(h)\,, \tag{52}$$

$$S_i = \sum h_i\,, \tag{53}$$

$$\sum E(h) = 1\,, \tag{54}$$

$$D^{(2h+1)}_{ij\ldots}[\boldsymbol{h}] = 0\,, \tag{55}$$

where the range of h is

$$0 \leq h \leq S$$

in *half-integral* steps. Note also that no $\boldsymbol{h}$ exists for $h = 0$.

7. – Structure of the sub-algebra.

As stated in the last section, the component of $\boldsymbol{h}$ can generate a sub-algebra. In this case, because

$$h^2 h_i = h(h+1)h_i$$

the multipole tensors *do* form a basis for the sub-algebra. The modulus is $E(h)$, as demonstrated by

$$Eh_i = h_i E = \frac{h^2}{h(h+1)} h_i = h_i .$$

The sub-algebra is finite; the $(2h+1)$'st multipole tensor vanishes. The order of the sub-algebra is easily computed as

$$\sum_0^{2h} (2\rho + 1) = (2h+1)^2 .$$

In terms of the algebra of $\boldsymbol{S}$, the equation

$$\sum E(h) = 1$$

indicates a decomposition of the modulus into a sum of orthogonal idempotents. In the theory of algebras, the question arises as to whether an idempotent is primitive. An idempotent, e, is not primitive if there exists another idempotent, u, such that

$$eu = ue = u .$$

A non-primitive idempotent can be written as the sum of two orthogonal idempotents, for

$$e = u + (e - u) ,$$

where

$$u^2 = u ,$$

$$(e-u)^2 = e^2 - 2ue + u^2 = e - u ,$$

$$u(e-u) = (e-u)u = ue - u^2 = 0 .$$

If $E(h)$ is not primitive, the modulus of the $\boldsymbol{S}$ algebra can be further decomposed. In terms of the sub-algebra of $\boldsymbol{h}$, the question of the primitiveness of $E(h)$ can be phrased as a question as to the decomposition of the modulus.

The sub-algebra is characterized by

$$D^{(2h+1)}_{ij\ldots} = 0 \,. \tag{56}$$

Chose now an a_μ, a real unit vector such as was used in the discussion of the spherical tensors. Multiply Eq. (56) by $a_i a_{j\ldots}$:

$$a_i a_{j\ldots} D^{(2h+1)}_{ij\ldots} = 0 \,. \tag{57}$$

Let

$$A^{(n)} = a_i a_j \ldots a_n D^{(n)}_{ij\ldots n} \,, \tag{58}$$

where n is any integer. Clearly, $A^{(n)}$ is a product of E and a polynomial of degree n in $\boldsymbol{a}\cdot\boldsymbol{h}$. For example,

$$A^{(1)} = \boldsymbol{a}\cdot\boldsymbol{h} = E\,\boldsymbol{a}\cdot\boldsymbol{h} \,,$$

$$A^{(2)} = a_i a_j (h_i h_j + h_j h_i - \tfrac{2}{3} h(h+1) E\,\delta_{ij}) \,,$$

$$= 2(\boldsymbol{a}\cdot\boldsymbol{h})^2 - \tfrac{2}{3} h(h+1) E \,,$$

$$= [2(\boldsymbol{a}\cdot\boldsymbol{h})^2 - \tfrac{2}{3} h(h+1)] E \,.$$

Eq. (57) is equivalent to

$$A^{(2h+1)} = 0 \,. \tag{59}$$

Here again, as in the last section, we have a Hamilton-Cayley equation. It is possible to factor Eq. (59) so as to read

$$\prod_{p=0}^{2h} (\boldsymbol{a}\cdot\boldsymbol{h} - \lambda_p) E = 0 \,,$$

where the λ_p are the as yet unknown roots of the Hamilton-Cayley equation. Formally, it is permissible to introduce

$$\varrho_h(m) = \prod_{\substack{p=0 \\ \neq m}}^{2h} \left(\frac{\boldsymbol{a}\cdot\boldsymbol{h} - \lambda_p}{\lambda_m - \lambda_p} \right) E(h) \,,$$

and it would follow that

$$(60)\qquad \begin{cases} \sum_m \varrho_h(m) = E(h) \\ \boldsymbol{a}\cdot\boldsymbol{h} = \sum \lambda_m \varrho_h(m) \\ \varrho_h(m)\,\varrho_h(m') = 0\,, & m \neq m' \\ \varrho_h(m)\,\varrho_h(m) = \varrho_h(m) \end{cases}$$

and

$$(61)\qquad \begin{cases} \boldsymbol{a}\cdot\boldsymbol{h}\,\varrho_h(m) = \varrho_h(m)\,\boldsymbol{a}\cdot\boldsymbol{h} = \lambda_m \varrho_h(m) \\ h^2 \varrho_h(m) = h(h+1)\,\varrho_h(m)\,. \end{cases}$$

To determine the λ_m, bring in the shift operators:

$$\boldsymbol{a}\cdot\boldsymbol{h}\,\boldsymbol{a}^{-}\cdot\boldsymbol{h}\,\varrho_h(m) = \boldsymbol{a}^{-}\cdot\boldsymbol{h}\,\boldsymbol{a}\cdot\boldsymbol{h}\,\varrho_h(m) - \boldsymbol{a}^{-}\cdot\boldsymbol{h}\,\varrho_h(m) = (\lambda_m - 1)\,\boldsymbol{a}^{-}\cdot\boldsymbol{h}\,\varrho_h(m)\,.$$

Multiply on the left by $\varrho_h(n)$:

$$(62)\qquad \varrho_h(n)\,\boldsymbol{a}\cdot\boldsymbol{h}\,\boldsymbol{a}^{-}\cdot\boldsymbol{h}\,\varrho_h(m) = (\lambda_m - 1)\,\varrho_h(n)\,\boldsymbol{a}^{-}\cdot\boldsymbol{h}\,\varrho_h(m)\,.$$

But $\boldsymbol{a}\cdot\boldsymbol{h}$ can also act to the left, so that the last equation is equal to

$$(63)\qquad \lambda_n\,\varrho_h(n)\,\boldsymbol{a}\cdot\boldsymbol{h}\,\varrho_h(m)\,.$$

Comparison of Eqs. (62) and (63) leads to

$$(64)\qquad \varrho_h(n)\,\boldsymbol{a}^{-}\cdot\boldsymbol{h}\,\varrho_h(m) = 0$$

unless

$$\lambda_n = \lambda_m - 1\,.$$

Eq. (64) can be summed over all values of n. Because

$$\sum \varrho_h(m) = E(h)$$

it follows that

$$\boldsymbol{a}^{-}\cdot\boldsymbol{h}\,\varrho_h(m) = 0$$

unless there exists a $\lambda = \lambda_m - 1$.

There are only $2\lambda+1$ roots. Let $\underline{\lambda}$ be the minimum root. Then

$$\boldsymbol{a}^- \cdot \boldsymbol{h}\boldsymbol{m}\varrho_h(\underline{\lambda}) = 0 .$$

Multiply by $\boldsymbol{a}^+ \cdot \boldsymbol{h}$:

$$\boldsymbol{a}^+\boldsymbol{h}\,\boldsymbol{a}^- \cdot \boldsymbol{h}\,\varrho_h(\lambda) = 0 .$$

Eq. (32) can now be used to give

$$[h^2 - \boldsymbol{a}\cdot\boldsymbol{h}(\boldsymbol{a}\cdot\boldsymbol{h} - 1)]\varrho_h(\underline{\lambda}) = 0$$

which requires

$$[h(h+1) - \underline{\lambda}(\underline{\lambda} - 1)]\varrho_h(\underline{\lambda}) = 0$$

and therefore

$$h(h+1) - \underline{\lambda}(\underline{\lambda} - 1) = 0 .$$

The two solutions of the equation are

$$\underline{\lambda} = h + 1$$

$$\underline{\lambda} = -h$$

and, since $\underline{\lambda}$ is the minimum root (h is positive),

(65) $$\underline{\lambda} = -h .$$

It is impossible that

$$\boldsymbol{a}^+ \cdot \boldsymbol{h}\,\varrho_h(\underline{\lambda}) = 0$$

for this would imply

$$\boldsymbol{a}^- \cdot \boldsymbol{h}\,\boldsymbol{a}^+ \cdot \boldsymbol{h}\,\varrho_h(\underline{\lambda}) = 0$$

or, by Eq. (32)

$$[h^2 - \boldsymbol{a}\cdot\boldsymbol{h}(\boldsymbol{a}\cdot\boldsymbol{h} + 1)]\varrho_h(\underline{\lambda}) = 0$$

or

$$h(h+1) - \underline{\lambda}(+\underline{\lambda}\,1) = 0$$

which is not possible unless $\underline{\lambda} = h = 0$, the case for which no $\boldsymbol{h}$ exists. Therefore

$$\boldsymbol{a}^+ \cdot \boldsymbol{h}\,\varrho_h(\underline{\lambda}) \neq 0 .$$

It can be proved by the methods already applied that

$$\boldsymbol{a}^+ \cdot \boldsymbol{h}\,\varrho_h(m) = 0$$

unless there exists a λ such that

$$\lambda = \lambda_m + 1 .$$

In particular, there must exist a $\lambda = \underline{\lambda} + 1$. In this way, all the roots are found to be the integers, or half-integers, going from

$$-h \leqslant \lambda_m \leqslant h . \tag{66}$$

At this point, the modulus of the $\boldsymbol{S}$ algebra is expressible as a sum of idempotents, $E(h)$, which characterize S^2, and these idempotents can be further resolved into idempotents, $\varrho_h(m)$, which characterize $\boldsymbol{a}\cdot\boldsymbol{S}$.

8. – The super algebra.

The product of any two elements of an algebra is an element of the algebra. In particular, the product of any two basal units can be expressed as a linear combination of the basal units. The scalar coefficients can be arranged into a multiplication table. In this section, it will be indicated how such a multiplication table could be constructed for an angular momentum algebra. However, it seems that for a large class of problems, it is not necessary to know each entry in the table. Only certain entries, or certain combinations of entries, are of physical significance. In any case, the constants of multiplication (*) can be specified when it is recognized that to every angular momentum algebra there is associated another algebra, a master or super algebra.

The super algebra is formed in this way: denote the algebra in which the angular momentum is imbedded as A. Consider the product of any two elements in A, xy. Represent the product as the result of an operation on y by x (from the left):

$$xy \Rightarrow x_{\text{op}} y$$

or, equally well:

$$xy \Rightarrow xy_{\text{op}}$$

(from the right). This set of operations, abstracted from the basic algebra, is itself an algebra, the super algebra. It is not a trivial algebra, for consider the operation of commutation with S_i,

$$g_i = (S_i, \ \ ) . \tag{67}$$

(*) L. E. Dickson (1), p. 17.

g_i is an angular momentum, for

$$\begin{aligned}(g_i, g_j)x &= (S_i, (S_j, x)) - (S_j, (S_i, x)) \\ &= (S_i, (S_j, x)) + (S_i, (x, S_j)) + (x, (S_j, S_i)) \\ &= (x, (S_j, S_i)) \\ &= i\varepsilon_{ij\sigma}(S_\sigma, x)\end{aligned}$$

or

$$(g_i, g_j) = i\varepsilon_{ij\sigma}g_\sigma \tag{68}$$

Note that

$$g^2 = (S_\mu, (S_\mu, \ )) \ . \tag{69}$$

Beyond its character as an angular momentum, g_i is distinguished as a distributive operation (*):

$$g_i(xy) = (g_i x)y + x(g_i y) \ . \tag{70}$$

Moreover, g_i can be written as the sum of two commuting angular momenta. Define the operations, m_i and n_i, by

$$m_i x = S_i x \tag{71}$$

$$n_i x = -xS_i \ . \tag{72}$$

Each of these operations is an angular momentum:

$$(m_i, m_j) = i\varepsilon_{ij\sigma}m_\sigma \tag{73}$$

$$(n_i, n_j) = i\varepsilon_{ij\sigma}n_\sigma \tag{74}$$

and they commute:

$$(m_i, n_j) = 0 \ . \tag{75}$$

Eqs. (73), (74), and (75) are easily verified. Clearly,

$$g_i = m_i + n_i \ . \tag{76}$$

(*) Sometimes called a derivation.

If the constraint on $\boldsymbol{S}$ is

$$D^{(2\mathcal{S}+1)}_{ij\ldots}[\boldsymbol{S}] = 0\,,$$

the constraints on $\boldsymbol{m}$ and $\boldsymbol{n}$ are

$$D^{(2\mathcal{S}+1)}_{ij\ldots}[\boldsymbol{m}] = 0\,, \tag{77}$$

$$D^{(2\mathcal{S}+1)}_{ij\ldots}[\boldsymbol{n}] = 0\,. \tag{78}$$

In Appendix IV, it is proved that Eqs. (76), (77) and (78) lead to

$$D^{(4\mathcal{S}+1)}_{ij\ldots}[\boldsymbol{g}] = 0\,. \tag{79}$$

Here is a set of angular momenta $\boldsymbol{g}$, $\boldsymbol{m}$, $\boldsymbol{n}$, to which the methods of this paper, and all other papers on angular momentum, can be applied. Even the vector space on which they operate is given; it is the algebra A. The modulus of the super algebra can be reduced into a sum of orthogonal, angular momentum idempotents, each of which is a product of commuting idempotents: a $(g^2, \boldsymbol{a}\cdot\boldsymbol{g})$ idempotent, an m^2 idempotent, and an n^2 idempotent. Each element of A can then be written, identically, as a sum of projections onto simultaneous eigenvectors of g^2, $\boldsymbol{a}\cdot\boldsymbol{g}$, m^2, n^2. Already certain eigenvectors have been identified. The spherical tensor

$$Y^{-N}_{N}[\boldsymbol{h}] = a^-_i \ldots a^-_N D^{(N)}_{ij\ldots}[\boldsymbol{h}]$$

satisfies the eigenvalue equations:

$$g^2 Y^{-N}_{N} = N(N+1)Y^{-N}_{N}$$

by Eq. (16), and

$$\begin{aligned}\boldsymbol{a}\cdot\boldsymbol{g}\,Y^{-N}_{N} &= a^-_i \ldots a^-_N a_\mu(S_\mu, D^{(N)}_{ij}\ldots) = i a^-_i \ldots a^-_N a_\mu \,\|\, \varepsilon_{\mu i\sigma} D^{(N)}_{\sigma j}\ldots \\ &= iN a^-_i \ldots a^-_N a_\mu \varepsilon_{\mu i\sigma} D^{(N)}_{\sigma j}\ldots = -N a^-_\sigma a^-_j \ldots a^-_N D^{(N)}_{\sigma j}\ldots = -N Y^{-N}_{N}\end{aligned}$$

and

$$m^2 Y^{-N}_{N}[\boldsymbol{h}] = h(h+1)Y^{-N}_{N}[\boldsymbol{h}]\,,$$

$$n^2\, Y^{-N}_{N}[\boldsymbol{h}] = h(h+1)Y^{-N}_{N}[\boldsymbol{h}]\,.$$

Y^{N-K}_{N}, which belongs to the eigenvalue $-N+K$ of $\boldsymbol{a}\cdot\boldsymbol{g}$, is generated by K operations with $\boldsymbol{a}^+\cdot\boldsymbol{g}$, the shift-up operator for $\boldsymbol{a}\cdot\boldsymbol{g}$.

The multiplication table of A could now be constructed by application of the combinatorial methods of angular momentum ([3,4]). To amplify this state-

ment, the case of an intrinsic angular momentum will be considered and only those constants of multiplication will be derived which are of physical significance.

9. – Calculus of an intrinsic angular momentum.

In this Section, S^2 is a constant, $\mathcal{S}(\mathcal{S}+1)$ times the unit element. Without going into the details of how to assure the equivalence (1), let the algebra of $\boldsymbol{S}$ be taken as equivalent to a matrix algebra of order $(2\mathcal{S}+1)$ (2). S_i is hermitian.

The primitive idempotents, $\varrho_{\mathcal{S}}(m)$, are the projection operators for $\boldsymbol{a}\cdot\boldsymbol{S}$. In a representation in which $\boldsymbol{a}\cdot\boldsymbol{S}$ is a diagonal matrix, $\varrho_{\mathcal{S}}(m)$ has unity as its m-th diagonal element and zeros elsewhere. In a representation in which $\boldsymbol{a}\cdot\boldsymbol{S}$ is not diagonal, a diagonal element of $\varrho_{\mathcal{S}}(m)$ is to be interpreted as the probability that the system be in the $\boldsymbol{a}\cdot\boldsymbol{S}=m$ state. The expectation value of any operators, X, in a state characterized by a definite value, m, of $\boldsymbol{a}\cdot\boldsymbol{S}$ is, (7):

$$\operatorname{Tr}\varrho_{\mathcal{S}}(m)X\,.$$

Therefore, only that part of the multiplication table which enters into the evaluation of traces of products need be investigated.

The algebra of $\boldsymbol{S}$ is the vector space of its super algebra. A metric can be introduced: on the matrix algebra generated by $\boldsymbol{S}$ define the inner product of x and y as

$$\operatorname{Tr} x^{\dagger}y = \langle x\,|\,y\rangle\,,$$

where $x^{\dagger}$ is the hermitian conjugate of x. This trace operation satisfies all the conditions of an inner product (8). With this definition of the inner product, it is easily proved that the linear operator's $\boldsymbol{g}$, $\boldsymbol{m}$, and $\boldsymbol{n}$ are all hermitian.

In a previous Section, the operator

$$A^{(n)} = a_i a_j \ldots a_n D^{(n)}_{ij\ldots n}$$

was introduced. For the case of a spin of multiplicity $2\mathcal{S}+1$, there are $2\mathcal{S}+1$ such operators $(0 \leqslant n \leqslant 2\mathcal{S})$. Each is a polynomial in $\boldsymbol{a}\cdot\boldsymbol{S}$. The set of $A^{(n)}$ is a set of commuting operators, a set of linearly independent commuting operators. The $A^{(n)}$ can be simultaneously diagonalized and, in diagonal form, the array of eigenvalues of an $A^{(n)}$, to be denoted as $A^{(n)}(m)$, can be considered

(7) P. A. M. Dirac: *The Principles of Quantum Mechanics* (Oxford, 1947).
(8) P. Halmos: *Finite Dimensional Vector Spaces* (Princeton, 1948), p. 126.
(9) M. Born: *Atomic Physics* (New York, 1946), p. 222.

a $2\mathcal{S}+1$ dimensional vector. The vectors, $A^{(n)}(m)$, span the space of $2\mathcal{S}+1$ dimensional vectors. In fact, the set of $A^{(n)}(m)$ is an orthogonal set. That is,

$$\sum_m A^{(n)}(m)A^{(n')}(m) = 0\,, \qquad n \neq n'. \tag{80}$$

For,

$$\sum_m A^{(n)}(m)A^{(n')}(n) = \operatorname{Tr} A^{(n)}A^{(n')} = \langle A^{(n)} | A^{(n')} \rangle$$

and $A^{(n)}$ and $A^{(n')}$ are each eigenvectors of g^2, but each belongs to a different eigenvalue:

$$g^2 A^{(n)} = n(n+1)A^{(n)}\,.$$

Eigenvectors of a hermitian operator, belonging to different eigenvalues, are orthogonal.

There is another way of proving this orthogonality relationship. The trace of a commutator is zero. Therefore, the trace of any multipole tensor, except the modulus, is zero:

$$\operatorname{Tr} D^{(n)}_{ij\ldots} = 0\,, \qquad n \neq 0. \tag{81}$$

because

$$n(n+1)D^{(n)}_{ij\ldots} = (\mathcal{S}_\mu, (\mathcal{S}_\mu, D^{(n)}_{ij\ldots}))\,.$$

What of the product of two multipole tensors? $D^{(n)}_{ij\ldots n}$ is an eigenvector of g^2 belonging to the eigenvalue $n(n+1)$; $D^{(n')}_{ij\ldots n'}$ is an eigenvector of g^2 belonging to the eigenvalue $n'(n'+1)$. Remember that g_i, an angular momentum, has a distributive property:

$$g_i(xy) = (g_i x)y + x(g_i y)\,.$$

Because of this special property of g_i, it is fruitful to think of the product of two multipole tensors as an outer product, an element of the direct product of the vector space with itself. The methods of angular momentum synthesis can be taken over. The product

$$D^{(n)}_{ij\ldots n} D^{(n')}_{i'j'\ldots n'}$$

presents a problem in the coupling of two angular momenta. If an angular momentum of $g^2 = n(n+1)$ is coupled to an angular momentum of $g^2 = n'(n'+1)$, then the resultant angular momentum can take on the values

$$n+n'\,, \qquad n+n'-1\,, \quad \ldots\,, \qquad |n-n'|\,.$$

Therefore, in the product, $D^{(n)}_{ij\ldots n} D^{(n')}_{i'j'\ldots n'}$, there will appear only those multipole of orders

$$n+n', \qquad n+n'-1, \qquad \ldots, \qquad |n-n'| .$$

If $n \neq n'$, the modulus does not appear, and so

$$\text{Tr}\, D^{(n)}_{ij\ldots n} D^{(n')}_{i'j'\ldots n'} = 0 , \qquad n \neq n'. \tag{82}$$

Consequently,

$$\text{Tr}\, A^{(n)} A^{(n')} = 0 , \qquad n \neq n'.$$

The $A^{(n)}$ are orthogonal. How are they normalized? The norm of $A^{(n)}$ is

$$\text{Tr}\,[A^{(n)}]^2 = \sum_m A^{(n)}(m) A^{(n)}(m) .$$

$A^{(n)}(m)$ is a polynomial in m. A recursion relation can be derived by multiplying Eq. (20), the recursion relation for $D^{(n)}_{ij\ldots n}$, by $a_i a_j \ldots a_n$:

$$A^{(n)} = n\boldsymbol{a}\cdot\boldsymbol{S}A^{(n-1)} - \frac{2(n-1)^2}{(2n-1)(2n-3)}\left[\mathcal{S}(\mathcal{S}+1) - \frac{n(n-2)}{4}\right]\frac{n(n-1)}{2} A^{(n-2)} ,$$

or

$$A^{(n)}(m) = nmA^{(n-1)}(m) - \frac{n(n-1)^3}{4(2n-1)(2n-3)} (2\mathcal{S}+n)(2\mathcal{S}-n+2) A^{(n-2)}(m) . \tag{83}$$

The equation indicates that the $A^{(n)}(m)$ are polynomials of definite and alternating parity.

Let

$$A^{(n)}(m) = g_n p_n(m) , \tag{84}$$

where $p_n(m)$ is to be an orthogonal polynomial normalized to unity:

$$\sum_m p_n(m) p_{n'}(m) = \delta_{nn'} . \tag{85}$$

Then the norm of $A^{(n)}(m)$,

$$\sum_m A^{(n)}(m) A^{(n)}(m) = g_n^2 . \tag{86}$$

There is a formula, called the Christoffel-Darboux formula, which holds for any orthonormalized polynomial [10] which, when applied to a polynomial

[10] G. Szegö: *Orthogonal Polynomials* (New York, 1939).

of alternating parity such as $p_n(m)$, reads as

$$p_n(m) = F_n m p_{n-1} - K_n p_{n-2} \,; \tag{87}$$

where

$$K_n = F_n / F_{n-1} \,.$$

Substitution of Eq. (84) into Eq. (83) gives a recursion relation for p_n. The Christoffel-Darboux formula is satisfied if

$$F_n^2 = \frac{4(4n^2-1)}{n^2(p^2-n^2)} \,, \tag{88}$$

and

$$g_n^2 = \frac{(n\,!)^4}{4^n[(2n-1)\,!\,!]^2} \cdot \frac{(p+n)\,!}{(2n+1)(p-n-1)\,!} \,, \tag{89}$$

where

$$p = 2\mathcal{S} + 1 \tag{90}$$

and

$$(2n-1)\,!\,! = (2n-1)(2n-3)(2n-5)\ldots \,.$$

For purposes of numerical computation, a generating function for p_n can be derived. This is done in Appendix V. The result is

$$p_n(q) = n\,! \left[\frac{(2n+1)(p-n-1)\,!}{(p+n)\,!}\right]^{\frac{1}{2}} \Delta^n \binom{q}{n} \binom{q-p}{n} \,, \tag{91}$$

with

$$m = -\mathcal{S} + q \,, \qquad 0 \leqslant q \leqslant 2\mathcal{S}$$

and Δ is the difference operator:

$$\Delta f(x) = f(x+1) - f(x)$$

and

$$\binom{q}{n} = \frac{q\,!}{n\,!\,(q-n)\,!} \,.$$

It is marvelous that these polynomials are already known and are associated with the name of the ubiquitous TCHEBICHEF (*).

(*) G. SZEGÖ [10], p. 32.

The initial definition of $A^{(n)}$ was

$$A^{(n)} = a_i a_j \dots a_n D^{(n)}_{ij\dots n} .$$

It is convenient to modify the scalar tensor, $a_i a_j \dots a_n$, so as to have a vanishing contraction. A set of scalar multipole tensors can be constructed out of a_i by the recursion relation

$$a^{(n)}_{ij\dots n} = \frac{1}{n}\left\{ a_i a^{(n-1)}_{j\dots} - \frac{2}{2n-1} a_\mu a^{(n-1)}_{\mu k\dots} \delta_{ij} \right\},$$

with $a^{(0)} = 1$, $a^{(1)}_i = a_i$. The scalar multipole tensors are to be fully symmetric in the indices and are to have vanishing contraction. By the methods used in the case of the angular momentum tensors, it can be proved that

$$a_\mu a^{(n)}_{\mu j\dots n} = \frac{n}{2n-1} a^{(n-1)}_{j\dots n} , \tag{92}$$

so that the equation for $a^{(n)}_{ij\dots n}$ reduces to

$$a^{(n)}_{ij\dots n} = \frac{1}{n}\left\{ a_i a^{(n-1)}_{j\dots} - \frac{2(n-1)}{(2n-1)(2n-3)} \delta_{ij} a^{(n-2)}_{k\dots n} \right\}. \tag{93}$$

If $b^{(n)}_{ij\dots n}$ is a scalar multipole tensor formed from the unit vector b_i, then

$$a^{(n)}_{ij\dots n} b^{(n)}_{ij\dots n} = \frac{n!}{(2n-1)!!} P_n(\boldsymbol{a}\cdot\boldsymbol{b}) , \tag{94}$$

where $P_n(x)$ is the ordinary Legendre polynomial normalized so that

$$P_n(1) = 1 .$$

Eq. (94) is derived in Appendix VI.

There will be a requirement, further on, to compute

$$\operatorname{Tr} A^{(n)} B^{(n)} ,$$

where

$$A^{(n)} = a^{(n)}_{ij\dots n} D^{(n)}_{ij\dots n} ,$$

$$B^{(n)} = b^{(n)}_{ij\dots n} D^{(n)}_{ij\dots n} .$$

The computation begins with

$$\operatorname{Tr} A^{(n)} B^{(n)} = a^{(n)}_{ij\dots n} b^{(n)}_{\bar{i}\bar{j}\dots\bar{n}} \operatorname{Tr} D^{(n)}_{ij\dots n} D^{(n)}_{\bar{i}\bar{j}} . \tag{95}$$

Now

$$\mathrm{Tr}\, D^{(n)}_{ij\ldots} D^{(n)}_{\bar{i}\bar{j}} = \lambda_n \underset{i\ldots}{\|} \delta_{i\bar{i}}\delta_{j\bar{j}} \ldots +$$

+ terms which yield zero when multiplied by multipole tensor coefficients.

Multiply by $a^{(n)}_{ij\ldots} a^{(n)}_{\bar{i}\bar{j}\ldots}$:

$$\mathrm{Tr}\,[A^{(n)}]^2 = \lambda_n n!\, a^{(n)}_{ij\ldots} a^{(n)}_{ij\ldots} = \lambda_n \frac{(n!)^2}{(2n-1)!!},$$

so that

$$\lambda_n = \frac{(2n-1)!!}{(n!)^2}\, g_n^2 . \tag{96}$$

Therefore, returning to Eq. (95),

$$\mathrm{Tr}\, A^{(n)} B^{(n)} = g_n^2 P_n(\boldsymbol{a}\cdot\boldsymbol{b}) . \tag{97}$$

For the computation of an expectation value, the factored forms of $\varrho_{\mathcal{S}}(m)$,

$$\varrho_{\mathcal{S}}(m) = \prod_{\substack{k=-\mathcal{S} \\ \neq m}}^{\mathcal{S}} \frac{\boldsymbol{a}\boldsymbol{S}\cdot - k}{m-k} ,$$

is not best manipulated. $\varrho_{\mathcal{S}}(m)$ is an operator which commutes with $\boldsymbol{a}\cdot\boldsymbol{S}$ and the $A^{(n)}$. As a polynomial in $\boldsymbol{a}\cdot\boldsymbol{S}$, $\varrho_{\mathcal{S}}(m)$ must be expandable into a sum of $A^{(n)}$:

$$\varrho_{\mathcal{S}}(m) = \sum C_k A^{(k)} . \tag{98}$$

To evaluate the coefficients, multiply Eq. (98) by $A^{(n)}$, and take the trace:

$$\mathrm{Tr}\, A^{(n)} \varrho_{\mathcal{S}}(m) = C_n \,\mathrm{Tr}\,[A^{(n)}]^2 .$$

But $\mathrm{Tr}\, A^{(n)} \varrho_{\mathcal{S}}(m)$ is the diagonal element of $A^{(n)}$:

$$A^{(n)}(m) = \mathrm{Tr}\, A^{(n)} \varrho_{\mathcal{S}}(m) = C_n \,\mathrm{Tr}\,[A^{(n)}]^2$$

so that

$$g_n p_n(m) = C_n g_n^2$$

or

$$C_n = \frac{p_n(m)}{g_n} .$$

The expression for the projection operator is

$$\varrho_{\mathcal{S}}(m) = \sum \frac{p_n(m)}{g_n} A^{(n)} . \tag{99}$$

10. – An illustrative problem.

Given a spin definitely quantized along a unit vector $\boldsymbol{a}$ with component m, what is the probability that it is quantized along a unit vector $\boldsymbol{b}$ with component m'? The answer is given, in effect, in the texts ([3]), ([4]). From the definitions given there, it follows that the probability is the square of the matrix element of the rotation matrix:

$$P_{mm'} = |\bar{a}^{\mathcal{S}}_{m'm}(\beta)|^2 ,$$

where β is the angle between the vectors $\boldsymbol{a}$ and $\boldsymbol{b}$. $\bar{a}^{\mathcal{S}}_{m'm}(\beta)$ is computed as

$$\bar{a}^{\mathcal{S}}_{m'm}(\beta) = [(\mathcal{S}+m)!\,(\mathcal{S}-m)!\,(\mathcal{S}+m')!\,(\mathcal{S}-m')!]^{\frac{1}{2}} \cdot \sum_r \frac{(-1)^r}{(\mathcal{S}-m'-r)!\,(\mathcal{S}+m-r)!\,(r+m'-m)!\,r!} \cdot \left(\cos\frac{\beta}{2}\right)^{2\mathcal{S}+m-m'-2r} \left(-\sin\frac{\beta}{2}\right)^{m'-m+2r} .$$

The calculus of the preceeding Section effects a straight-forward answer to the problem and a more compact expression:

$$P_{mm'} = \mathrm{Tr}\, \varrho(m)\varrho'(m') ,$$

with

$$\varrho(m) = \sum \frac{p_n(m)}{g_n} A^{(n)} ,$$

$$\varrho'(m') = \sum \frac{p_n(m')}{g} B^{(n)} .$$

By the formulae of the preceding Section:

$$P_{mm'} = \sum p_n(m) p_n(m') P_n(\boldsymbol{a}\cdot\boldsymbol{b}) . \tag{100}$$

11. – The Tchebichef polynomials and Clebsch-Gordan coefficients.

There is a startling identification of the Tchebichef polynomials, $p_n(\mathcal{S}; m)$ with certain Clebsch-Gordan coefficients. It is a fact that

$$p_n(\mathcal{S}; m) = (-1)^{\mathcal{S}-m} C(\mathcal{S}, \mathcal{S}, n; m, -m) , \tag{101}$$

where the C-notation is that of ROSE [1]. To establish this identity, examine the relationships of the eigenvectors of the super algebra:

Because

$$\boldsymbol{a}\cdot\boldsymbol{m}\,\varrho_{\mathcal{S}}(m) = \quad m\varrho_{\mathcal{S}}(m)$$

$$\boldsymbol{a}\cdot\boldsymbol{n}\;\varrho_{\mathcal{S}}(m) = -\,m\varrho_{\mathcal{S}}(m)$$

$$m^2\varrho_{\mathcal{S}}(m) = n^2\varrho_{\mathcal{S}}(m) = \mathcal{S}(\mathcal{S}+1)\varrho_{\mathcal{S}}(m)$$

$$\mathrm{Tr}\;\varrho_{\mathcal{S}}(m)\varrho_{\mathcal{S}}(m) = \langle\varrho_{\mathcal{S}}(m)\,|\,\varrho_{\mathcal{S}}(m)\rangle = 1$$

the projection operator is, except for a phase factor, a normalized, simultaneous eigenvector of m^2, n^2, $\boldsymbol{a}\cdot\boldsymbol{m}$, and $\boldsymbol{a}\cdot\boldsymbol{n}$.

Because

$$\boldsymbol{a}\cdot\boldsymbol{g}\,\frac{A^{(n)}}{g_n} = 0\,,$$

$$g^2\,\frac{A^{(n)}}{g_n} = n(n+1)\,\frac{A^{(n)}}{g_n}\,,$$

$$m^2\,\frac{A^{(n)}}{g_n} = n^2\,\frac{A^{(n)}}{g_n} = \mathcal{S}(\mathcal{S}+1)\,\frac{A^{(n)}}{g_n}\,,$$

$$\mathrm{Tr}\left[\frac{A^{(n)}}{g_n}\right]^2 = \left\langle\frac{A^{(n)}}{g_n}\,\middle|\,\frac{A^{(n)}}{g_n}\right\rangle = 1\,,$$

$A^{(n)}/g^n$ is, except for a phase factor, a normalized eigenvector of g^2, $\boldsymbol{a}\cdot\boldsymbol{g}$, m^2, and n^2. The equations

$$\varrho\mathcal{S}(m) = \sum p_n(\mathcal{S};\,m)\,\frac{A^{(n)}}{g_n}\,,$$

$$\frac{A^{(n)}}{g_n} = \sum p_n(\mathcal{S};\,m)\,\varrho\mathcal{S}(m)\,,$$

are expressions of the transformation from the uncoupled representation to the coupled representation and inverse transformation. The Clebsch-Gordan coefficients are defined as the elements of the unitary transformation. $p_n(\mathcal{S}, m)$ is, therefore, proportional to $C(\mathcal{S}, \mathcal{S}, n;\, m, -m)$. Comparison of formulae gives the phase factor in Eq. (101).

Appendix I

Reduction of a symmetric tensor.

Given a tensor, $t_{ij\ldots n}$, symmetric in all its indices, it is required to work upon it so as to assure a vanishing contraction.

Consider a term of the following sort: select p pairs of indices and contract $t_{ij\ldots n}$ over each pair, ending up with a p-times contracted tensor such as $t_{\mu\mu\nu\nu\ldots mn}$. (There are p dummy indices). Now distribute the selected $2p$ indices over p Kronecker deltas. Form the product of the Kronecker deltas and $t_{\mu\mu\nu\nu\ldots mn}$. Repeat the process so as to bring the form to full symmetry. Call this symmetrized form, made of p-times contracted tensors and Kronecker deltas, $T^{(p)}_{ij\ldots n}$.

The number of terms in $T^{(p)}_{ij\ldots n}$ is

$$\frac{n!}{(n-2p)!}\frac{1}{p!}\frac{1}{2^p},$$

for, $n!/(n-2p)!$ is the number of ways of leaving behind $(n-2p)$ equivalent indices.

$1/p!$ is to take care of the permutation of the Kronecker symbols.

$1/2^p$ accounts for the symmetry of each Kronecker delta (*).

The modified tensor shall be

$$t_{ij\ldots n} + c_1 T^{(1)}_{ij\ldots n} + \ldots + c_k T^{(k)}_{ij\ldots n} + \cdots .$$

Its contraction shall vanish. Under the contraction of i and j, let's say, $T^{(p)}_{ij\ldots n}$ splits into two types of terms. There were the terms in which the pair, (i, j), appeared in the $t_{\mu\mu ij\ldots}$ factor. These $(n-2)!/[(n-2-2p)!\,p!\,2^p]$ terms end up containing $(p+1)$-times contracted tensors. The second type of terms in the contracted $T^{(p)}_{ij\ldots n}$ are

$$\frac{n!}{(n-2p)!\,p!\,2^p} - \frac{(n-2)!}{(n-2-2p)!\,p!\,2^p} = \frac{(n-2)!\,(2n-2p-1)}{(n-2p)!\,(p-1)!\,2^{p-1}},$$

in number, and these second type terms retain a p-times contracted tensor. The second type terms in the contracted $T^{(p)}$ duplicate the first type terms in the contracted $T^{(p-1)}$. There is over duplication, since the ratio of the

---

(*) $n!/((n-2p)!\,p!\,2^p)$ is an integer, since

$$\frac{n!}{(n-2p)!\,p!\,2^p} = \frac{n!}{(n-2p)!\,(2p)!}\left\{\frac{(2p)!}{(2p-2)!\,2!}\,\frac{(2p-2)!}{(2p-4)!\,2!}\cdots\right\},$$

a product of integers which suggests another combinatorial synthesis.

number of $T^{(p)}$ second type terms to the number of $T^{(p-1)}$ first type terms is

$$\frac{(n-2)!\,(2n-2p-1)}{(n-2p)!\,(p-1)!\,2^{p-1}} \Big/ \frac{(n-2)!}{(n-2-2p+2)!\,(p-1)!\,2^{p-1}} = 2n-2p-1\,.$$

It must be remembered when i and j appeared in different Kronecker deltas in $T^{(p)}$, the contraction gave only one contribution. When i and j appeared in the same Kronecker delta, the contraction gave three contributions. The over duplication is therefore

$$(2n-2p-1)+2 = 2n-2p+1\,.$$

To assure a vanishing contraction chose

$$c_{p-1}/c_p = -[2(n-p)+1]\,,$$

or

$$c_p = -\left\{\frac{1}{2(n-p)+1}\right\} c_{p-1}\,.$$

Appendix II

The linear independence of multipole tensor components.

Given that

$$\sum_{(n)} a^{(n)}_{ij\ldots n} D^{(n)}_{ij\ldots n} = 0\,,$$

where $a^{(n)}_{ij\ldots n}$ is a field quantity, symmetric in all its indices and of vanishing contraction, to prove that

$$a^{(n)}_{ij\ldots n} = 0\,.$$

To begin with, it is no restriction to have $a^{(n)}_{ij\ldots n}$ fully symmetric and of vanishing contraction. Care must be taken in the application of this theorem to alter the coefficients so that the conditions are satisfied.

Repeated application of Eq. (16),

$$(S_\mu, (S_\mu, D^{(n)}_{ij\ldots n})) = n(n+1) D^{(n)}_{ij\ldots n}\,,$$

will give rise to an unending set of simultaneous equations of the form

$$\sum_{(n)} n^k (n+1)^k a^{(n)}_{ij\ldots n} D^{(n)}_{ij\ldots n} = 0\,.$$

There is no linear dependence of the equations since the coefficients, $n^k(n+1)^k$, in the k-th equation, being polynomials of $2k$ degree in n, are not dependent on the polynomial coefficients of the other equations. Therefore, it must be that

$$a^{(n)}_{ij\ldots n} D^{(n)}_{ij\ldots n} = 0 \,. \tag{AII.1}$$

The idea is to somehow get from (AII.1) to

$$a^{(n)}_{\mu\nu\ldots} D^{(n)}_{ij\ldots} = 0 \,, \tag{AII.2}$$

which is a simple product with no summation involved. To get from (AII.1) to (AII.2) one must, so to speak, release the indices. The inner product form in (AII.1) can be said to have zero indices released; the form in (AII.2) can be said to have n indices released. An intermediate form such as

$$a^{(n)}_{\mu\nu k\ldots n} D^{(n)}_{ijk\ldots n} \,,$$

can be said to have two indices released. Given that the form with p indices released is null, does it follow that the form with $p+1$ indices released is null? It does.

Let

$$a^{(n)}_{\mu\nu\ldots m\ldots n} D^{(n)}_{ij\ldots m\ldots n} = 0 \,, \tag{AII.3}$$

where p indices are released. Commute the equation with S_λ:

$$a^{(n)}_{\mu\nu\ldots m\ldots n} \Big\| _{ij} \varepsilon_{\lambda i\sigma} D^{(n)}_{\sigma j\ldots m\ldots n} = 0 \,.$$

To the right of the double bar there will be a number of tensors in which p indices are released. Therefore, by (AII.3), the equation can be reduced to

$$a^{(n)}_{\mu\nu\ldots m\ldots n} \Big\| _{m\ldots n} \varepsilon_{\lambda m\sigma} D^{(n)}_{ij\ldots\sigma\ldots n} = 0 \,.$$

Because of the symmetry of $a^{(n)}_{\mu\nu\ldots m\ldots n}$, this last equation can be reduced further to

$$\varepsilon_{\lambda m\sigma} a^{(n)}_{\mu\nu\ldots m\ldots n} D^{(n)}_{ij\ldots\sigma\ldots n} = 0 \,.$$

There's no need to encumber the typography with the released indices. The last equation can be written as

$$\varepsilon_{\lambda m\sigma} a^{(n)}_{x,m\ldots n} D^{(n)}_{\bar{x},m\ldots n} = 0 \,, \tag{AII.4}$$

where the single index to the left of the comma represents the set of released

indices. The necessary and sufficient condition that (AII.4) be true is

$$a^{(n)}_{x,m\ldots n} D^{(n)}_{\bar{x},\sigma\ldots n} = a^{(n)}_{x,\sigma\ldots n} D^{(n)}_{\bar{x},m\ldots n} \,. \tag{AII.5}$$

Commute both sides of (AII.5) with S_λ, and multiply by $\varepsilon_{\nu\lambda\sigma}$:

$$a^{(n)}_{x,m\ldots n}\varepsilon_{\nu\lambda\sigma}(S_\lambda, D^{(n)}_{\bar{x}\,\sigma\ldots n}) = a^{(n)}_{x,\sigma\ldots n}\varepsilon_{\nu\lambda\sigma}(S_\lambda, D^{(n)}_{\bar{x},m\ldots n}) \,,$$

$$(n+1)a^{(n)}_{x.m\ldots n} D^{(n)}_{\bar{x},\nu\ldots n} = a^{(n)}_{x,\sigma\ldots n}\varepsilon_{\nu\lambda\sigma} \underset{m\ldots n}{\|} \varepsilon_{\lambda m\gamma} D^{(n)}_{\bar{x},\gamma\ldots n} + a^{(n)}_{x,\sigma\ldots n}\varepsilon_{\nu\lambda\sigma} \underset{\bar{x}}{\|} \varepsilon_{\lambda\bar{x}\gamma} D^{(n)}_{\gamma,m\ldots n} =$$

$$= a^{(n)}_{x,m\ldots n} D^{(n)}_{\bar{x},\nu\ldots n} - a^{(n)}_{x,\sigma\ldots n} D^{(n)}_{\bar{x},\sigma\ldots n}\,\delta_{m\nu} + [n-(p+1)]a^{(n)}_{x,\sigma\sigma\ldots} D^{(n)}_{\bar{x},m\nu\ldots} -$$

$$- [n-(p+1)]a^{(n)}_{x,\sigma\nu\ldots} D^{(n)}_{\bar{x},m\sigma\ldots} + \underset{\bar{x}}{\|} a^{(n)}_{x,x\ldots n} D^{(n)}_{\nu,m\ldots n} - \underset{\bar{x}}{\|} a^{(n)}_{x,\nu\ldots n} D^{(n)}_{\sigma,m\ldots n}\,\delta_{\nu\bar{x}} \,.$$

The second term on the right hand side of the equation can be dropped because of (AII.3). The third term can be dropped because $a^{(n)}_{x,m\ldots n}$ has vanishing contraction. The sixth term can be written, because of the $D^{(n)}_{\sigma,m\ldots n}$ symmetry, as

$$\underset{\bar{x}}{\|} a^{(n)}_{x,\sigma\ldots n} D^{(n)}_{m,\sigma\ldots n}\,\delta_{\nu\bar{x}} \,,$$

which vanishes because of (AII.3). All that is left is

$$(n+1)a^{(n)}_{x,m\ldots n} D^{(n)}_{\bar{x},\nu\ldots n} = a^{(n)}_{x,m\ldots n} D^{(n)}_{\bar{x},\nu\ldots n} - [n-(p+1)]a^{(n)}_{x,\nu\sigma\ldots} D^{(n)}_{\bar{x},n\sigma\ldots} + \tag{AII.6}$$

$$+ \underset{\bar{x}}{\|} a^{(n)}_{x,\bar{x}\ldots n} D^{(n)}_{\nu,m\ldots} \,.$$

The last term on the right hand side of this equation can be re-arranged by virtue of (AII.5):

$$\underset{\bar{x}}{\|} a^{(n)}_{x,\bar{x}\ldots} D^{(n)}_{\nu,m\ldots} = \underset{\bar{x}}{\|} a^{(n)}_{x,m\ldots} D^{(n)}_{\nu,\bar{x}\ldots} = \underset{\bar{x}}{\|} a^{(n)}_{x,m\ldots} D^{(n)}_{\bar{x},\nu\ldots} = p a^{(n)}_{x,m\ldots} D^{(n)}_{\bar{x},\nu\ldots} \,.$$

Application of (AII.5) to the second term of (AII.6) allows a final combination into

$$[2n-(2p+1)]a^{(n)}_{x,m\ldots} D^{(n)}_{\bar{x},\nu\ldots} = 0 \,.$$

It is impossible that $2n-(2p+1)=0$. Therefore,

$$a^{(n)}_{x,m\ldots} D^{(n)}_{\bar{x},\nu\ldots} = 0 \,,$$

which is the statement that the form with $p+1$ released indices vanishes.

There is a route from

$$a^{(n)}_{ij\ldots} D^{(n)}_{ij\ldots} = 0 \,,$$

to

$$a^{(n)}_{\mu\nu\ldots} D^{(n)}_{ij\ldots} = 0 \,.$$

$a^{(n)}_{\mu\nu\ldots}$ is a member of the field. Unless

$$D^{(n)}_{ij\ldots} = 0 ,$$

it must be that

$$a^{(n)}_{\mu\nu\ldots} = 0 .$$

Appendix III

The number of independent multipole tensor components.

Let $t_{ij\ldots n}$ be any tensor symmetric in all its indices. There need be no connection with angular momentum. The number of indices is n; the range of each index is x, not necessarily equal to three. To count the number of non-redundant elements in $t_{ij\ldots n}$, list the elements in dictionary order. That is, order the elements $t_{ij\ldots n}$ so that $i \leqslant j \leqslant k \leqslant \ldots \leqslant n$. This is an exhaustive, non-redundant procedure.

It is possible to partition the dictionary into disjoint classes whose sum is the dictionary. Two elements of the dictionary are termed equivalent if they have the same number of indices set equal to unity, the lowest value in the range. This equivalence relation is reflexive, symmetrical, and transitive. The class (p) is composed of all those elements in which p and only p indices are equal to unity. Every element of the dictionary must be in one class and one class only.

Denote the total number of elements in the dictionary by $f(x, n)$. The class (p) has $f(x-1, n-p)$ elements, for it reads as a sub-dictionary based on $n-p$ indices each with a range of $x-1$. Therefore,

$$f(x, n) = \sum_{p=0}^{n} f(x-1, n-p) .$$

By a change of summation index, the equation becomes

$$f(x, n) = \sum_{p=0}^{n} f(x-1, p) , \tag{AIII.1}$$

from which a recursion relation is deriveable as

$$f(x, n) = f(x, n-1) + f(x-1, n) . \tag{AIII.2}$$

It is necessary to set

$$f(x, 0) = 1 .$$

All scalars are linearly dependent. With this boundary condition (AIII.1) is deriveable from (AIII.2). The solution is

$$f(x, n) = \frac{(n+x-1)!}{n!\,(x-1)!} .$$

This formula is the Bose-Einstein combinatorial formula. Let n be the number of particles in a particular one-particle energy state, x be the degeneracy of that state. Then $f(x, n)$ is the number of distinguishable arrangements of particles among the degenerate levels belonging to that energy state [9].

To get the number of independent elements of $D^{(n)}_{ij\ldots n}$, account must be taken of the dependence brought in by the conditions of vanishing contraction. There are $\frac{1}{2}n(n-1)$ such conditions. The number of independent elements of $D^{(n)}_{ij\ldots n}$ is, therefore,

$$f(3, n) - \tfrac{1}{2}n(n-1) = 2n+1 .$$

Appendix IV

Proof that $D^{(4\mathcal{S}+1)}_{ij\ldots}[g] = 0$.

Given that

$$D^{(2\mathcal{S}+1)}_{ij\ldots}[\boldsymbol{m}] = 0 ,$$

and

$$D^{(2\mathcal{S}+1)}_{ij\ldots}[\boldsymbol{n}] = 0 ,$$

to prove that

$$D^{(4\mathcal{S}+1)}_{ij\ldots}[\boldsymbol{g}] = 0 ,$$

where $\boldsymbol{g} = \boldsymbol{m} + \boldsymbol{n}$. The understanding is that no lesser multipole tensor of $\boldsymbol{m}$ or $\boldsymbol{n}$ vanishes, and that no lesser multipole tensor of $\boldsymbol{g}$ will vanish.

The proof is carried out in the spherical representation. As $\boldsymbol{g} = \boldsymbol{m} + \boldsymbol{n}$, it follows that

$$\text{(AIV.1)} \qquad \begin{cases} Y^1_1[\boldsymbol{g}] = Y^1_1[\boldsymbol{m}] + Y^1_1[\boldsymbol{n}], \\ Y^p_p[\boldsymbol{g}] = \sum_{k=0}^{p} \dfrac{p!}{k!\,(p-k)!}\, Y^{p-k}_{p-k}[\boldsymbol{m}]\, Y^k_k[\boldsymbol{n}], \end{cases}$$

because $\boldsymbol{m}$ and $\boldsymbol{n}$ commute. In particular,

$$Y^{4\mathcal{S}}_{4\mathcal{S}}[\boldsymbol{g}] = \sum_{k=0}^{4\mathcal{S}} \frac{(4\mathcal{S})!}{k!\,(4\mathcal{S}-k)!}\, Y^{4\mathcal{S}-k}_{4\mathcal{S}-k}[\boldsymbol{m}]\, Y^k_k[\boldsymbol{n}] \propto Y^{2\mathcal{S}}_{2\mathcal{S}}[\boldsymbol{m}]\, Y^{2\mathcal{S}}_{2\mathcal{S}}[\boldsymbol{n}],$$

because $Y^k_k[\boldsymbol{n}] = 0$ for $k > 2\mathcal{S}$, and $Y^{4\mathcal{S}-k}_{4\mathcal{S}-k}[\boldsymbol{m}] = 0$ for $k \leqslant 2\mathcal{S}$.

Therefore, $Y^{4\mathcal{S}}_{4\mathcal{S}}[\boldsymbol{g}] \neq 0$. However,

$$\begin{aligned} Y^{4\mathcal{S}+1}_{4\mathcal{S}+1}[\boldsymbol{g}] &= \sum_{k=0}^{4\mathcal{S}+1} \frac{(4\mathcal{S}+1)!}{k!\,(4\mathcal{S}+1-k)!}\, Y^{4\mathcal{S}+1-k}_{4\mathcal{S}+1-k}[\boldsymbol{m}]\, Y^k_k[\boldsymbol{n}] \\ &= \sum_{k=0}^{2\mathcal{S}} \frac{(4\mathcal{S}+1)!}{k!\,(4\mathcal{S}+1-k)!}\, Y^{4\mathcal{S}+1-k}_{4\mathcal{S}+1-k}[\boldsymbol{m}]\, Y^k_k[\boldsymbol{n}] \\ &= \sum_{k=1}^{2\mathcal{S}+1} \frac{(4\mathcal{S}+1)!}{(2\mathcal{S}+1-k)!\,(2\mathcal{S}+k)!}\, Y^{2\mathcal{S}+k}_{2\mathcal{S}+1}[\boldsymbol{m}]\, Y^{2\mathcal{S}+1-k}_{2\mathcal{S}+1-k}[\boldsymbol{n}] \\ &= 0 . \end{aligned}$$

which is equivalent to

$$D^{(4\mathcal{S}+1)}_{ij\ldots}[\boldsymbol{g}] = 0 .$$

516

Appendix V

The Tchebichef polynomials.

The Tchebichef polynomial, $p_n(m)$, is given by the diagonal matrix element of $A^{(n)}$ in a representation in which $\boldsymbol{a}\cdot\boldsymbol{S}$ is diagonal:

$$p_n(m) = \langle m | \frac{A^{(n)}}{g_n} | m \rangle .$$

$A^{(n)}$ must be proportional to the spherical tensor Y_n^0 which is generated by n commutations of $\boldsymbol{a}^+\cdot\boldsymbol{S}$ with the spherical tensor Y_n^{-n} defined as

$$Y_n^{-n} = (\boldsymbol{a}^-\cdot\boldsymbol{S})^n .$$

That is,

$$A^{(n)} = R(\boldsymbol{a}^+\cdot\boldsymbol{S}, (\boldsymbol{a}^+\cdot\boldsymbol{S}), \ldots (\boldsymbol{a}^+\cdot\boldsymbol{S}), (\boldsymbol{a}^-\cdot\boldsymbol{S})^n) ,$$

in which n commutations are performed. This can be written as

$$A^{(n)} = R(\boldsymbol{a}^+\cdot\boldsymbol{S}^{[n]}, (\boldsymbol{a}^-\cdot\boldsymbol{S})^n) .$$

In general,

$$(X^{[n]}, Y) = \sum_{k=0}^{n} (-1)^k \frac{n!}{k!\,(n-k)!} X^{n-k} Y X^k ,$$

so that

$$p_n(m) = \frac{R}{g_n} \sum_{k=0}^{n} (-1)^k \frac{n!}{k!\,(n-k)!} \langle m | (\boldsymbol{a}^+\cdot\boldsymbol{S})^{n-k} (\boldsymbol{a}^-\cdot\boldsymbol{S})^n (\boldsymbol{a}^+\cdot\boldsymbol{S})^k | m \rangle .$$

The matrix product is reduced to a single term because of the shift nature of the operators

$$p_n(m) = \frac{R}{g_n} \sum_{k=0}^{n} (-1)^k \frac{n!}{k!\,(n-k)!} \langle m | (a^+\cdot S)^{n-k} | m-n+k \rangle \cdot \\ \cdot \langle m-n+k | (\boldsymbol{a}^-\cdot\boldsymbol{S})^n | m+k \rangle \langle m+k | (\boldsymbol{a}^+\cdot\boldsymbol{S})^k | m \rangle .$$

The matrix elements of the shift operators, given in many texts, are

$$\langle m+1 | \boldsymbol{a}^+\cdot\boldsymbol{S} | m \rangle = c_m ,$$
$$\langle m-1 | \boldsymbol{a}^-\cdot\boldsymbol{S} | m \rangle = c_{m-1}^* ,$$

where

$$|c_m|^2 = \tfrac{1}{2}[\mathcal{S}(\mathcal{S}+1) - m(m+1)] .$$

Therefore,

$$p_n(m) = \frac{R}{g_n} \sum_{k=0}^{n} (-1)^k \frac{n!}{k!\,(n-k)!} \prod_{1}^{n-k} |c_{m-p}|^2 \prod_{0}^{k-1} |c_{m+p}|^2 .$$

The products can be evaluated:

$$\prod_{1}^{n-k} |c_{m-p}|^2 = \frac{1}{2^{n-k}} \frac{(\mathcal{S}-m+n-k)!}{(\mathcal{S}-m)!} \frac{(\mathcal{S}+m)!}{(\mathcal{S}+m-n+k)!},$$

$$\prod_{0}^{k-1} |c_{m+p}|^2 = \frac{1}{2^k} \frac{(\mathcal{S}-m)!}{(\mathcal{S}-m-k)!} \frac{(\mathcal{S}+m+k)!}{(\mathcal{S}+m)!} .$$

Going on:

$$p_n(m) = \frac{1}{2^n} \frac{R}{g_n} \sum_{k=0}^{n} (-1)^k \frac{n!}{k!\,(n-k)!} \frac{(\mathcal{S}-m+n-k)!}{(\mathcal{S}-m-k)!} \frac{(\mathcal{S}+m+k)!}{(\mathcal{S}+m+k-n)!} .$$

Let $m = -\mathcal{S} + q$, so that

$$p_n(q) = \frac{1}{2^n} \frac{R}{g_n} \sum_{k=0}^{n} (-1)^k \frac{n!}{k!\,(n-k)!} \frac{(2\mathcal{S}-q+n-k)!}{(2\mathcal{S}-q-k)!} \frac{(q+k)!}{(q+k-n)!} .$$

Let

$$t(q) = \frac{(2\mathcal{S}-q+n)!}{n!\,(2\mathcal{S}-q)!} \frac{q!}{n!\,(q-n)!},$$

so that

$$p_n(q) = \frac{1}{2^n} \frac{R}{g_n} (n!)^2 \sum_{k=0}^{n} (-1)^k \frac{n!}{k!\,(n-k)!} t(q+\varkappa) =$$

$$= \frac{1}{2^n} \frac{R}{g_n} (n!)^2 \sum_{k=0}^{n} (-1)^k \frac{n!}{k!\,(n-k)!} \exp\left[k \frac{\mathrm{d}}{\mathrm{d}q}\right] t(q) =$$

$$= (-1)^n \frac{1}{2^n} \frac{R}{g_n} (n!)^2 \left(\exp\left[\frac{\mathrm{d}}{\mathrm{d}q}\right] - 1\right)^n t(q) = (-1)^n \frac{1}{2^n} \frac{R}{g_n} (n!)^2 \Delta^n t(q) ,$$

where

$$\Delta f(x) = \left(\exp\left[\frac{\mathrm{d}}{\mathrm{d}q}\right] - 1\right) f(x) = f(x+1) - f(x) .$$

To simplify the form of $t(q)$, note the identity

$$\frac{(2\mathcal{S}-q+n)!}{(2\mathcal{S}-q)!} = (-1)^n \frac{(q-2\mathcal{S}-1)!}{(q-2\mathcal{S}-1-n)!},$$

which permits

$$p_n(q) = \frac{1}{2^n} \frac{R}{g_n} (n!)^2 \Delta^n \binom{q}{n} \binom{q-p}{n},$$

with $p = 2\mathcal{S}+1$.

To evaluate R, realize that the coefficient of $(\boldsymbol{a}\cdot\boldsymbol{S})^n$ in $A^{(n)}$ is $n!$, and that, consequently, the coefficient of q^n in $p_n(q)$ is $n!/g_n$. Now,

$$\Delta^n f(x) = \frac{\mathrm{d}^n}{\mathrm{d}x^n} f(x+\theta n), \qquad 0 \leqslant \theta \leqslant 1,$$

by the mean-value theorem quoted by SZEGÖ [10]. Therefore, the coefficient of q^n in $(-1)^n t(q)$ is

$$\frac{2n(2n-1)\ldots(n+1)}{(n!)^2} = \frac{2^n(2n-1)!!}{(n!)^2}.$$

It must be that

$$R = \frac{n!}{(2n-1)!!}.$$

With the choice of g_n as always positive

$$p_n(q) = n!\left[\frac{(2n+1)(p-n-1)!}{(p+n)!}\right]^{\frac{1}{2}} \Delta^n \binom{q}{n}\binom{q-p}{n}.$$

APPENDIX VI

Evaluation of $a^{(n)}_{ij\ldots n}\, b^{(n)}_{i\ldots n}$.

The recursion relation for $a^{(n)}_{ij\ldots n}$ is

$$a^{(n)}_{ij\ldots n} = \frac{1}{n}\left\{ a_i a^{(n-1)}_{j\ldots n} - \frac{2(n-1)}{(2n-1)(2n-3)}\, \delta_{ij} a^{(n-2)}_{k\ldots n} \right\}.$$

Multiply by $b_i b_j \ldots b_n$

$$b_i b_j \ldots b_n a^{(n)}_{ij\ldots n} = \frac{1}{n}\left\{ n\boldsymbol{a}\cdot\boldsymbol{b}\, b_j \ldots b_n a^{(n-1)}_{j\ldots} - \frac{2(n-1)}{(2n-1)(2n-3)}\,\frac{n(n-1)}{2}\, b_k \ldots b_n a^{(n-2)}_{k\ldots n} \right\}.$$

Let

$$Z_n = b_i b_j \ldots b_n a^{(n)}_{ij\ldots n} = b^{(n)}_{ij,\ldots n} a^{(n)}_{ij\ldots n}.$$

The recursion relation for Z_n is the last equation:

$$Z_n = \boldsymbol{a}\cdot\boldsymbol{b}\, Z_{n-1} - \frac{(n-1)^2}{(2n-1)(2n-3)}\, Z_{n-2}, \tag{AIV.1}$$

with $Z_1 = \boldsymbol{a}\cdot\boldsymbol{b}$.

The Legendre polynomials, $P_n(x)$, are completely defined by the recursion relation [11]

$$xP_{n-1} = \frac{n}{2n-1} P_n + \frac{n-1}{2n-1} P_{n-2} ,$$

with the condition $P_1(x) = x$. It follows from (AVI.1) that

$$Z_n = a^{(n)}_{ij\ldots n} b^{(n)}_{ij\ldots n} = \frac{n!}{(2n-1)!!} P_n(\boldsymbol{a}\cdot\boldsymbol{b}) .$$

[11] E. JAHNKE - F. EMDE: *Tables of Higher Functions* (Leipzig, 1952), p. 111.

SUPPLEMENTO AL VOLUME XII, SERIE X
DEL NUOVO CIMENTO

N. 1, 1959
2° Trimestre

Ultrasonic Investigation of Nuclear Spin-Lattice Relaxation.

G. S. VERMA

Department of Physics, University of Allahabad - Allahabad,
Ultrasonics Research Laboratory, Western Reserve University - Cleveland, Ohio

(ricevuto il 17 Dicembre 1958)

CONTENTS. — 1. Introduction. – 2. Theory. – 3. Experimental technique. – 4. Results and discussion.

1. – Introduction.

Ultrasonics has been utilised extensively in the past to study the properties of gases and liquids which threw light on many molecular processes and has proved to be a very convenient tool in the study of relaxation phenomena. In solids it has been used very effectively in studying the elastic constants from velocity measurements, especially in single crystals. From damping and attenuation measurements there have been attempts to seek information regarding imperfections- dislocations and radiation damage and their kinetics. Recently Ultrasonics has proved to be very helpful in studying the properties of conduction electrons in metals, especially when the metal becomes superconducting (1). The number of thermally excited electrons can be determined from the measurements of ultrasonic absorption in superconductors. As a matter of fact, MORSE and BOHM (2) claim to have measured the energy gap in the electron energy spectrum in superconducting tin and indium, which, however, has been questioned by some workers (3). Ultra-

(1) W. P. MASON and H. E. BOMMEL: *Journ. Acoust. Soc. Am.*, **28**, 930 (1956).
(2) R. W. MORSE and A. V. BOHM: *Phys. Rev.*, **108**, 1094 (1958).
(3) M. A. BIONDI, A. T. FORRESTER, M. P. GARFUNKEL and C. B. SATTERTHWAITE: *Rev. Mod. Phys.*, **30**, 1109 (1958).

sonic waves have been shown to induce transitions ([4,5]) in the spin energy levels of a particular nucleus through the nuclear spin-lattice relaxation, the possibility of which was suggested first of all by ALTSCHULER ([6]). The possibility of a strong interaction between ultrasonic waves and spin waves, especially when both the magnon and phonon frequencies and wavelengths are equal, has been discussed in ferromagnetic crystals and shown to provide an independent measure of the exchange energy constant ([7]). The potentialities of Ultrasonics in the Gigahertz region, where the acoustic frequencies have wavelengths comparable to the wavelength of light and the phonon energy may approach thermal energy at liquid helium temperatures, are many but its application is, however, limited by the inherent experimental difficulties. The highest frequency which has been achieved as reported recently by BARANSKII ([8]), and BOMMEL and DRANSFELD ([9]) is only 2.5 GHz ([*]).

In many solids nuclear spin lattice relaxation is known to involve the presence of impurities. It has been shown by ROLLIN and HATTON ([10,11]) and by BLOEMBERGEN ([12]) that for nuclei with a spin $I=\frac{1}{2}$, which have no quadrupole moments, the observed spin-lattice relaxation times can be satisfactorily accounted for by assuming that they are due to the presence of electronically paramagnetic impurities. However, for nuclei with a spin $I>\frac{1}{2}$, POUND ([13]) has obtained strong evidence that in pure crystals at not too low temperatures the quadrupolar relaxation may be more important than the impurity relaxation. KRANENDONK ([14]) has developed a method for calculating the relaxation time due to the perturbation of this quadrupolar interaction by the lattice vibrations. The interaction between atomic nuclei and their sorroundings can be studied by nuclear paramagnetic resonance ([15]) or nuclear induction ([16]). Such interactions affect relaxation times ([17,18]), line widths ([17,19]) and frequen-

([4]) W. G. PROCTOR and W. TANTTILA: *Phys. Rev.*, **98**, 1854 (1955).
([5]) W. G. PROCTOR and W. TANTTILA: *Phys. Rev.*, **101**, 1757 (1956).
([6]) S. A. ALTSCHULER: *Dokl. Akad. Nauk USSR*, **85**, 1235 (1952).
([7]) C. KITTEL: *Phys. Rev.*, **110**, 836 (1958).
([8]) K. N. BARANSKIJ: *Dokl. Akad. Nauk SSSR*, **114**, 517 (1957).
([9]) H. E. BOMMEL and K. DRANSFELD: *Phys. Rev. Lett.*, **1**, 234 (1958).
([*]) *Note added in proof.* – Recently JACOBSEN has extended the frequency to the vicinity of 10 GHz (*Phys. Rev. Lett.*, **2**, 249 (1959)).
([10]) J. HATTON and B. V. ROLLIN: *Proc. Roy. Soc.*, A **199**, 222 (1949).
([11]) N. BLOEMBERGEN: *Physica*, **15**, 386 (1946).
([12]) B. V. ROLLIN and J. HATTON: *Phys. Rev.*, **74**, 346 (1948).
([13]) R. V. POUND: *Phys. Rev.*, **79**, 685 (1950).
([14]) J. VAN KRANENDONK: *Physica*, **20**, 781 (1954).
([15]) E. M. PURCELL, H. C. TORREY and R. V. POUND: *Phys. Rev.*, **69**, 37 (1946).
([16]) F. BLOCH, W. W. HANSEN and M. E. PACKARD: *Phys. Rev.*, **69**, 127 (1946).
([17]) N. BLOEMBERGEN, E. M. PURCELL and R. V. POUND: *Phys. Rev.*, **73**, 679 (1948).
([18]) N. BLOEMBERGEN: *Thesis* (Leiden, 1948).
([19]) E. M. PURCELL, N. BLOEMBERGEN and R. V. POUND: *Phys. Rev.*, **70**, 988 (1946).

522

cies ([20,21]). Excitation of ultrasonic vibrations in a crystal offers another method to study the strength and nature of the coupling between lattice vibrations and the nuclear quadrupole moment. Ultrasonic waves can induce nuclear spin transition in solids through the interaction of the quadrupole moment of the nucleus, with the electric field gradient at the site of the nucleus, which varies periodically with the frequency of sound waves.

2. – Theory.

The electrostatic interaction energy $\mathcal{H}$ of a nucleus with the charges constituting its environment is given by

$$\mathcal{H} = \int \varrho(x)\, V(x)\, \mathrm{d}^3x\,, \tag{1}$$

where the integration is taken over the nuclear volume, $\varrho(x)$ is the nuclear charge density and $V(x)$, the electrostatic potential arising from all charges other than those of the nucleus under consideration. Expanding the potential $V(x)$ in a power series about the nuclear center of mass, we obtain

$$\mathcal{H} = \int \mathrm{d}^3x\, \varrho(x) \left\{ V_0 + \sum_j \left(\frac{\partial V}{\partial x_j}\right)_0 x_j + \frac{1}{2} \sum_{j,k} \left(\frac{\partial^2 V}{\partial x_j \partial x_k}\right) x_j x_k + \dots \right\} =$$

$$= ZeV_0 + \sum_j P_j \left(\frac{\partial V}{\partial x_j}\right)_0 + \frac{1}{2} \sum_{j,k} Q_{jk} \left(\frac{\partial^2 V}{\partial x_j \partial x_k}\right)_0 + \dots\,,$$

where

$$\int \mathrm{d}^3x\, \varrho(x) \quad = Ze = \text{nuclear charge}\,,$$

$$\int \mathrm{d}^3x\, \varrho(x) x_j \quad = P_j = \text{electric dipole moment}\,,$$

$$\int \mathrm{d}^3x\, \varrho(x) x_j x_k = Q_{jk} = \text{electric quadrupole moment tensor}\,.$$

The first term represents the electrostatic energy of a point nucleus, which is independent of nuclear size, shape or orientation and is of no interest to us. The second term which arises due to electric dipole moment vanishes by virtue of the fact that the nuclear ground state is non-degenerate and the time

([20]) W. D. Knight: *Phys. Rev.*, **76**, 1259 (1949).

([21]) W. G. Proctor and F. C. Yu: *Phys. Rev.*, **77**, 717 (1950); W. C. Dickinson: *Phys. Rev.*, **77**, 736 (1950).

averaged charge distribution satisfies the symmetry condition $\varrho(x) = \varrho(-x)$.

Leaving all other non-vanishing terms in the series except the quadrupole terms, the interaction energy is represented by

$$\mathcal{H} = \tfrac{1}{2} \sum_{j,k} Q_{jk} V_{jk} , \tag{3}$$

where $V_{jk} = (\partial^2 V / \partial x_j \partial x_k)$ and Q_{jk} and V_{jk} are second rank symmetric tensors. V_{jk}, however, can be expressed as $V_{jk} = -\partial E_k / \partial x_j$, where $E_k = -\partial V / \partial x_k$ is the electric field at the nucleus. Thus V_{jk} represents the gradient of the electric field and interacts with the nuclear electric quadrupole moment.

However, in order to have correct evaluation of the interaction energy, one has to consider the polarization of the charge cloud surrounding the nucleus by nuclear quadrupole moment, which has been discussed in detail by STERNHEIMER and his associates ([22-24]). The induced quadrupole moment can be much larger than the nuclear quadrupole moment and is such as to reinforce the nuclear quadrupole moment. This is called antishielding or polarization effect. FOLEY, STERNHEIMER and TYCKO ([25]) find that in ions such as Cl^-, Cu^+, Rb^+ and Cs^+ the induced quadrupole moment is 10 to 100 times larger than the coupling obtained when there is no polarization effect. KRANENDONK has taken this effect into consideration by means of a multiplication factor

$$\mathcal{H}' = \alpha \mathcal{H} = \alpha \int \varrho(x)\, V(x)\, \mathrm{d}^3 x , \tag{4}$$

where $\mathcal{H}'$ is the interaction energy when polarization of the atomic core is taken into consideration. There is another effect which is also present. The direct field from the sorrounding charges polarizes the central ion and leads to an extra field in a direction opposite to the direct field. It is called the shielding effect and its effect on the quadrupolar spin-lattice relaxation is small.

KRANENDONK also drew attention to the effects of covalency which have been discussed in detail by YOSIDA and MORIYA ([26]). If the lattice is deformed by the lattice vibrations, mixing of p- or d-like orbitals into the s-like wavefunctions of ions takes place and as a result the electrons in these non-spherically symmetric orbitals give rise to an electric field gradient at the nucleus which interacts with nuclear quadrupole moments. The resulting quadrupole

([22]) R. M. STERNHEIMER: *Phys. Rev.*, **80**, 102 (1950).
([23]) R. M. STERNHEIMER: *Phys. Rev.*, **84**, 244 (1951).
([24]) R. M. STERNHEIMER and H. M. FOLEY: *Phys. Rev.*, **92**, 1460 (1953).
([25]) H. M. FOLEY, R. M. STERNHEIMER and D. TYCKO: *Phys. Rev.*, **93**, 734 (1954).
([26]) K. YOSIDA and T. MORIYA: *Journ. Phys. Soc. Japan*, **11**, 33 (1956).

521

spin-lattice coupling can be much larger than with the direct field alone. As a matter of fact YOSIDA and MORIYA have shown that the relaxation time due to covalency is shorter by a factor of 10^{-4} than in the pure ionic case.

A detailed theory of quadrupolar nuclear spin-lattice relaxation has been formulated by KRANENDONK for crystals of the NaCl-type. He considered one parameter model where the nearest neighbours of a given nucleus are replaced with point charges of magnitude $q = \gamma e$. The magnitude q of these charges is the only adjustable parameter of this model. The dimensionless quantity γ is a measure of the strength of the spin-lattice coupling and e is the electronic charge. Values of γ are obtained by comparing the experimental values of the relaxation time with the values calculated on the basis of the point charge model. For example, in ionic crystals of the NaCl-type spin lattice coupling is due to interaction of the nuclear quadrupole moment with the electric field arising from six equal point charges on the six neighbouring lattice sites and the rest of the charges are supposed to be absent from the crystal. According to his calculations, values of γ are of the order $10^2 \div 10^3$.

KRANENDONK expressed the interaction Hamiltonian as

$$\mathcal{H}' = \sum_{\mu} W_{\mu} Q_{\mu} , \tag{5}$$

where Q_{μ} are the components of the nuclear quadrupole moment, the equivalent operators of which in the subspace $I = \text{constant}$ are given by

$$\begin{cases} Q_0 = C[3I_z^2 - I(I+1)] , \\ Q_{\pm 1} = C[I_z(I_x \pm iI_y) + (I_x \pm iI_y)I_z]/2 , \\ Q_{\pm 2} = C[I_x \pm iI_y]^2 , \end{cases} \tag{6}$$

where

$$C = \frac{eQ}{I(2I-1)} , \qquad eQ = \langle II|Q_0|II\rangle .$$

The only non-vanishing matrix elements of Q_{μ} are $\langle m+\mu|Q_{\mu}|m\rangle$. Here mh and $(m+\mu)h$ represent the eigenvalues of I_z if the direction of the external applied magnetic field is along the z-axis. The coefficients of W_{μ} are functions of the displacement of the nuclei relative to the central nucleus and hence W_{μ} can be developed in powers of r_i.

$$W_{\mu} = A_{\mu} + \sum_i \boldsymbol{A}_{\mu i} \cdot \boldsymbol{r}_i + \sum_{ij} \boldsymbol{A}_{\mu ij} : \boldsymbol{r}_i \boldsymbol{r}_j + \cdots . \tag{7}$$

The double dot indicates the complete inner product.

Substituting (7) into equation (5), $\mathcal{H}'$ is expressed in powers of the r_i

$$\mathcal{H}' = \mathcal{H}'_0 + \mathcal{H}'_1 + \mathcal{H}'_2 + \dots , \tag{8}$$

where $\mathcal{H}'_0 = \sum_{\mu} A_\mu Q_\mu$.

The constant term $\mathcal{H}'_0$ gives rise to the static quadrupolar interaction. This perturbs the originally equally spaced Zeeman levels of the nucleus in the presence of the external magnetic field and leads to splitting of the magnetic resonance line. In the absence of external field only pure quadrupole spectrum occurs. However, relaxation is caused by the higher order terms $\mathcal{H}'_1$, $\mathcal{H}'_2$, ... The linear term $\mathcal{H}'_1$ in the first order leads to the so-called direct processes in which the nuclear spin makes an upward (downward) transition and one of the lattice oscillators is de-excited (excited). In second order the linear term $\mathcal{H}'_1$ gives rise to Raman processes where the nuclear transition is accompanied by the simultaneous excitation of one lattice oscillator and de-excitation of another one. The first order effect of $\mathcal{H}'_2$ gives rise to Raman processes. The contribution due to first and higher order effects of $\mathcal{H}'_1$, higher order effects of $\mathcal{H}'_2$, and first order effects of $\mathcal{H}'_3$, $\mathcal{H}'_4$, ... are all negligible compared to the first order effect of the quadratic term in $\mathcal{H}'_2$, which according to KRANENDONK, is the relevant relaxation process. Processes in which more than two phonons are involved are important only at very high temperatures.

The transition probability $P(m, m+\mu)$ that the nuclear spin makes a transition from the state m to the state $(m+\mu)$ as a result of the Raman processes is given by

$$P(m, m+\mu) = \frac{|Q_{\mu m}|^2}{2\pi^3 d^2 v^2} \int_0^{\omega_m} \frac{k^4 |f(\omega)|^2}{\omega^2} \frac{e^x}{(e^x-1)^2} M_\mu(ka)\, \mathrm{d}\omega , \tag{9}$$

where $Q_{\mu m} = \langle m+\mu | Q_\mu | m \rangle$, d = mass density of the crystal, v = velocity of propagation of long wavelength sound waves independent of direction of polarization, $k = 1/\omega$ the wave number of lattice waves. $\varrho(\omega)$ is defined such that the number of lattice oscillators with frequency between ω and $\omega+\mathrm{d}\omega$ is given by

$$\varrho(\omega)\, \mathrm{d}\omega = \frac{V}{2\pi^2 v} f(\omega) k^2\, \mathrm{d}\omega , \tag{10}$$

where V represent the volume of the crystal.

Here $x = \hbar\omega/\mathcal{K}T$,

$$M_\mu(ka) = \sum_{ss'} \sum_{ijlm} N_{\mu ijlm} L_{sil;\, s'jm}(ka) , \tag{11}$$

with

$$(12)\quad L_{sil;\,s'jm}(ka) = \{B_i(\boldsymbol{k}, s)B_l(\boldsymbol{k}, s)\}\{B_j(\boldsymbol{k}', s')B_m(\boldsymbol{k}', s')\}, \qquad N_{\mu ijlm} = \boldsymbol{A}_{\mu ej} : \boldsymbol{A}_{\mu lm} .$$

Here the $B_i(\boldsymbol{k}, s)$ are defined by

$$(13)\qquad B_i = B_i(\boldsymbol{k}, s) = \delta(s, 1)(\cos \boldsymbol{k}_i a_i - 1) + \delta(s, 2) \sin \boldsymbol{k}_i a_i ,$$

where s is a number which is equal to 1 or 2 and $a_i = \boldsymbol{R}_i - \boldsymbol{R}_0$, $\boldsymbol{R}_i$ and $\boldsymbol{R}_0$ representing the equilibrium positions of the ith and central nuclei, respectively; $\boldsymbol{k}$ and $\boldsymbol{k}'$ represent the directions of wave propagation. The curly brackets in the equation (12) represent averaging over all directions of $\boldsymbol{k}$ and $\boldsymbol{k}'$.

Das, Roy and Ghosh Roy ([27]) extended Kranendonk's method to body centered cubic structure.

Yosida and Moriya have calculated nuclear relaxation times in ionic crystals on the basis of covalent effects alone. In the covalent state one p-electron is transferred from the closed shell of the halogen ion to one of the neighbouring metal ions and the hole created by the missing electron causes an asymmetric electron distribution which results in a very large field gradient at the site of the halogen nucleus. They introduced a parameter λ which is a measure of the degree of covalency. Kanda ([28]), as matter of fact, has calculated the degree of covalency in each halide from his measurements of chemical shift of halogen ions in several metal halides and shown that the reciprocal of the relaxation time is proportional to the degree of covalency.

Taking as the zeroth approximation the purely ionic state Yosida and Moriya have calculated relaxation times on the basis of perturbation treatment. According to them the third order perturbation is given by

$$(14)\qquad \Delta E_3 = \sum_{n,n'} \frac{(g\,|\,\mathcal{H}_t\,|\,n)(n\,|\,\mathcal{H}_Q\,|\,n')(n'\,|\,\mathcal{H}_t\,|\,g)}{(E_n - E_g)(E_{n'} - E_g)} ,$$

where g means the ground state and n, n' are the excited covalent states in which one p-electron on the halogen ion is transferred to the neighbouring n or n'-th metal ion. Here $\mathcal{H}_t$ is the Hamiltonian which causes an electron of the halogen to jump to a neighbouring metal ion. Here terms such as $(n\,|\,\mathcal{H}_Q\,|\,n')$ are usually very small unless $n = n'$. Neglecting such terms

$$(15)\qquad \Delta E_3 = \sum_{n} \frac{\{(g\,|\,\mathcal{H}\,|\,n\}^2}{(E_n - E_g)^2}\,(n\,|\,\mathcal{H}_Q\,|\,n) = \sum_{n} \lambda_n (n\,|\,\mathcal{H}_Q\,|\,n) = \sum_{\mu} Q_\mu \sum_{n} \lambda_n\,(n\,|\,W_\mu\,|\,n) ,$$

([27]) T. P. Das, D. K. Roy and S. K. Ghosh Roy: *Phys. Rev.*, **104**, 1568 (1956).
([28]) T. Kanda: *Journ. Phys. Soc. Japan*, **10**, 85 (1955).

where $\lambda = |(g|\mathcal{H}_t|n)|^2/(E_n - E_g)^2$ is a measure of the degree of covalency and is regarded as a function of the distance only between the halogen nucleus and the n-th metal nucleus.

The second summation in equation (15), abbreviated as W_μ, is expressed as follows

$$(16)\qquad \begin{cases} W_0 = A' \sum_n \lambda_n (n_n^2 - \tfrac{1}{3}) , \\ W_{\pm 1} = 2A' \sum_n \lambda_n n_n (l_n \mp i m_n) , \\ W_{\pm 2} = \tfrac{1}{2} A' \sum_n \lambda_n (l_n \mp i m_n)^2 , \end{cases}$$

where l_n, m_n and n_n represent the direction cosines of $\boldsymbol{R}_n(X_n, Y_n, Z_n)$, which denotes the co-ordinates of n-th metal nucleus and $A' = \frac{3}{10} eC\langle 1/r^3\rangle_H$.

Expanding these expressions in powers of small deviations from equilibrium positions and retaining second order terms only

$$(17)\qquad \begin{cases} W_{\pm 1} = \dfrac{2A'}{a^2} \sum_n \big[u \{ n^\mp \gamma_n (\delta \boldsymbol{R}_n)^2 + 2(\boldsymbol{n}\cdot\delta\boldsymbol{R}_n)(\gamma_n \delta \boldsymbol{R}_n^\mp + n^\mp \delta Z_n)\} + \\ \qquad + v n^\mp \gamma_n (\boldsymbol{n}\cdot\delta\boldsymbol{R}_n)^2 + \lambda \delta Z_n \delta \boldsymbol{R}_n^\mp \big] , \\ W_{\pm 2} = \dfrac{A'}{2a^2} \sum_n \big[u \{ (n^\mp)^2 (\delta \boldsymbol{R}_n)^2 + 4 n^\mp \delta \boldsymbol{R}_n^\mp (\boldsymbol{n}\cdot\delta\boldsymbol{R}_n)\} + \\ \qquad + v (n^\mp)^2 (\boldsymbol{n}\cdot\delta\boldsymbol{R}_n)^2 + \lambda (n^\mp)^2 (\delta \boldsymbol{R}_n^\mp)^2 \big] , \end{cases}$$

where $u = \frac{1}{2} a\lambda' - \lambda$, $v = 4\lambda - \frac{5}{2} a\lambda' + \frac{1}{2} a^2 \lambda''$.

α_n, β_n and γ_n are the direction cosines of the equilibrium position of the n-th metal ion measured from the halogen nucleus, $\delta\boldsymbol{R}_n(\delta X_n, \delta Y_n, \delta Z_n)$ the displacement vector, $a = |\boldsymbol{R}_{n0}|$, $\boldsymbol{n}$ is the unit vector whose components are α_n, β_n and γ_n, $\delta\boldsymbol{R}_n^\pm = \delta X_n \pm i\,\delta Y_n$ and $\alpha_n \pm i\beta_n = n^\pm$.

Following Kranendonk's notation, the transition probability is given by

$$(18)\qquad P(m, m+\mu) = \frac{A'^2 |Q_{\mu m}|^2}{2\pi^3 a^4 d^2 v^2} \int_0^{\omega_m} \frac{k^4}{\omega^2} \frac{e^x}{(e^x - 1)^2} M_\mu(ka)\, \mathrm{d}\omega ,$$

where

$$(19)\qquad M_\mu(ka) = \sum_{n,n'} \sum_s \{ B_{n\sigma} B_{n\sigma'} B_{n'\sigma} B_{n'\sigma'} \} \sum_{pp'} \omega^*_{\mu n}(\sigma\sigma') \omega_{\mu n'}(\sigma\sigma') ,$$

where σ stands for a set of $\boldsymbol{k}$ (wave factor), p (one of the three directions of polarization) and s, s taking values 1 and 2,

$$B_n(\sigma) = \delta(s, 1)(\cos a\,\boldsymbol{k}\cdot n - 1) + \delta(s, 2) \sin a\,\boldsymbol{k}\cdot n$$

and

$$
\left\{
\begin{aligned}
&\sum_{pp'} \omega^*_{1n}(\sigma\sigma')\omega_{1n'}(\sigma\sigma') = 4\,[u^2\{15(\alpha_n\alpha_{n'}+\beta_n\beta_{n'})\gamma_n\gamma_{n'}+2(\alpha_n\alpha_{n'}+\beta_n\beta_{n'})(\boldsymbol{n}\cdot\boldsymbol{n}')+\\
&\quad+4(\alpha_n^2+\beta_n^2)\gamma_{n'}^2+4\gamma_n\gamma_{n'}(\boldsymbol{n}\cdot\boldsymbol{n}')\}+v^2\{\alpha_n\alpha_{n'}+\beta_n\beta_{n'})\gamma_n\gamma_{n'}(\boldsymbol{n}\cdot\boldsymbol{n}')^2\}+\\
&\quad+\lambda^2+uv\,\{2(\alpha_n\alpha_{n'}+\beta_n\beta_{n'})\gamma_n\gamma_{n'}+4(\alpha_n\alpha_{n'}+\beta_n\beta_{n'})\gamma_n^2(\boldsymbol{n}\cdot\boldsymbol{n}')+\\
&\quad+4(\alpha_n^2+\beta_n^2)\gamma_n\gamma_{n'}(\boldsymbol{n}\cdot\boldsymbol{n}')\}+\lambda u\,\{2(\alpha_n^2+\beta_n^2)+4\gamma_n^2\}+2\lambda v(\alpha_n^2+\beta_n^2)\gamma_n^2]\,,\\
&\sum_{pp'} \omega^*_{2n}(\sigma\sigma')\omega_{2n'}(\sigma\sigma') = u^2\,[\tfrac{11}{4}\{-(\alpha_n^2+\beta_n^2)(\alpha_{n'}^2+\beta_{n'}^2)+2(\alpha_n\alpha_{n'}+\beta_n\beta_{n'})^2\}+\\
&\quad+4(\alpha_n\alpha_{n'}+\beta_n\beta_{n'})(\boldsymbol{n}\cdot\boldsymbol{n}')+2(\alpha_n^2+\beta_n^2)(\alpha_{n'}^2+\beta_{n'}^2)]+\tfrac{1}{4}v^2\{-(\alpha_n^2+\beta_n^2)(\alpha_{n'}^2+\beta_{n'}^2)+\\
&\quad+2(\alpha_n\alpha_{n'}+\beta_n\beta_{n'})^2\}(\boldsymbol{n}\cdot\boldsymbol{n}')^2+\lambda^2+uv\,\{\tfrac{1}{2}(\alpha_n-i\beta_n)^2(\alpha_{n'}+i\beta_{n'})^2+\\
&\quad+2(\alpha_n\alpha_{n'}+\beta_n\beta_{n'})(\alpha_n^2+\beta_n^2)(\boldsymbol{n}\cdot\boldsymbol{n}')\}+4\lambda u(\alpha_n^2+\beta_n^2)+\tfrac{1}{2}\lambda v(\alpha_n^2+\beta_n^2)^2\,,
\end{aligned}
\right. \tag{20}
$$

which can be simplified further as $(\boldsymbol{n}\cdot\boldsymbol{n}')=0$ for $n\neq\pm n'$ and $(\boldsymbol{n}\cdot\boldsymbol{n}')=\pm 1$ for $n=\pm n'$ crystals of the NaCl type. The transition probability, however, is given in the following convenient form

$$
P(m,\, m+\mu) = C_{\mu m}\, T^{*2} \sum_{n=1}^{4} N_{\mu n} D_n(T^*)\,, \tag{21}
$$

where

$$
\begin{aligned}
C_{\mu m} &= 3\,|Q_{\mu m}|^2/\pi d^2 v^3 a^3 \quad \text{(Kranendonk)}\,,\\
&= A'^2 C^3 |Q_{\mu m}|^2/2\pi^3 a^7 d^2 v^3 \quad \text{(Yosida and Moriya)}\\
T^* &= T/\theta\,, \qquad \theta \text{ Debye temperature}\,,\\
D_n(T^*) &= T^* \int_0^{1/T^*} \frac{x^2 e^x}{(e^x-1)^2}\, L_n(cT^*x)\,\mathrm{d}x\,,
\end{aligned}
$$

$$
c = \underset{\text{for NaCl-type}}{k_m a = (6\pi^2)^{\frac{1}{3}}}\,, \qquad \hbar\omega_m = \hbar v k_m = k\theta\,, \qquad \underset{\text{for CsCl-type}}{c = (q\sqrt{3}\,\pi^2/2)^{\frac{1}{3}}}
$$

$D_n(T^*)$ has been expanded in descending powers of T^* by Kranendonk and is given as follows (for NaCl or simple cubic crystals)

$$
\left\{
\begin{aligned}
D_1(T^*) &= 0.20 - 0.007\,1/T^{*2} + \cdots\\
D_2(T^*) &= 0.92 - 0.055/T^{*2} + \cdots\\
D_3(T^*) &= 0.50 - 0.031/T^{*2} + \cdots\\
D_4(T^*) &= 0.28 - 0.017/T^{*2} + \cdots
\end{aligned}
\right. \tag{22}
$$

The coefficients $N_{\mu n}$ are given by

$$N_{\mu 1} = \sum (\pm) N_{\mu i i l m} ; \qquad N_{\mu n} = \sum N_{\mu i j l m} ; \qquad (n = 2, 3, 4),$$

$$N_{\mu i j l m} = \boldsymbol{A}_{\mu i j} \cdot \boldsymbol{A}_{\mu l m} ,$$

$$\boldsymbol{A}_{\mu i j} = 0 \quad \text{for } i \neq j , \qquad \boldsymbol{A}_{-\mu i i} = \boldsymbol{A}^{*}_{\mu i i} , \qquad \boldsymbol{A}_{\mu - i - i} = \boldsymbol{A}_{\mu i i} ,$$

$$A_{0ii} = A \begin{vmatrix} a_i & 0 & 0 \\ 0 & b_i & 0 \\ 0 & 0 & c_i \end{vmatrix} ; \qquad A_{1ii} = A \begin{vmatrix} 0 & 0 & c'_i \\ 0 & 0 & f'_i \\ e'_i & f'_i & 0 \end{vmatrix} ; \qquad A_{2ii} = A \begin{vmatrix} a''_i & d''_i & 0 \\ d''_i & b''_i & 0 \\ 0 & 0 & c''_i \end{vmatrix} ;$$

$$A = q/16a^5 ,$$

where a is the equilibrium distance between the nearest neighbours. According to Yosida and Moriya, the coefficients $N_{\mu n}$ are given by

$$(23) \qquad \begin{cases} N_{\mu 1} = N_{\mu 2} = \sum\limits_{\left(\substack{n,n' \\ n = \pm n'}\right)} \sum\limits_{\nu\nu'} \omega^{*}_{\mu n}(\sigma\sigma')\omega_{\mu n'}(\sigma\sigma') , \\ N_{\mu 3} = N_{\mu 4} = \sum\limits_{\left(\substack{n,n' \\ n \neq \pm n'}\right)} \sum\limits_{\nu\nu'} \omega^{*}_{\mu n}(\sigma\sigma')\omega_{\mu n'}(\sigma\sigma') , \end{cases}$$

which after somewhat lengthy calculations are expressed as functions of λ, λ' and λ''

$$(24) \qquad \begin{cases} N_{11} = N_{12} = 48\lambda^2 \left[\left\{ 1 - \frac{4}{3}\left(\frac{a\lambda'}{\lambda}\right) + \frac{2}{3}\left(\frac{a\lambda'}{\lambda}\right)^2 \right\} + \right. \\ \qquad \left. + 2\alpha \left\{ 1 + \frac{1}{3}\left(\frac{a\lambda'}{\lambda}\right) - \frac{1}{2}\left(\frac{a\lambda'}{\lambda}\right)^2 + \frac{1}{12}\left(\frac{a^2\lambda''}{\lambda}\right)^2 \right\} \right] , \\ N_{14} = 32\lambda^2 \left[\left\{ -1 + \left(\frac{a\lambda'}{\lambda}\right)^2 \right\} + \right. \\ \qquad \left. + 5\alpha \left\{ 1 - \frac{9}{20}\left(\frac{a\lambda'}{\lambda}\right)^2 + \frac{3}{5}\left(\frac{a^2\lambda''}{\lambda}\right) - \frac{1}{10}\left(\frac{a^2\lambda''}{\lambda}\right)\left(\frac{a\lambda'}{\lambda}\right) \right\} \right] , \\ N_{21} = N_{22} = 18\lambda^2 \left[\left\{ 1 - \frac{7}{9}\left(\frac{a\lambda'}{\lambda}\right) + \frac{5}{18}\left(\frac{a\lambda'}{\lambda}\right)^2 + \frac{1}{30}\left(\frac{a^2\lambda''}{\lambda}\right)^2 \right\} - \right. \\ \qquad \left. - \frac{1}{3}\alpha \left\{ 1 + \frac{1}{3}\left(\frac{a\lambda'}{\lambda}\right) - \frac{1}{2}\left(\frac{a\lambda'}{\lambda}\right)^2 + \frac{1}{12}\left(\frac{a^2\lambda''}{\lambda}\right)^2 \right\} \right] , \\ N_{24} = 2\lambda^2 \left[\left\{ 1 + \frac{7}{4}\left(\frac{a\lambda'}{\lambda}\right)^2 + 3\left(\frac{a^2\lambda''}{\lambda}\right) - \frac{1}{2}\left(\frac{a^2\lambda''}{\lambda}\right)\left(\frac{a\lambda'}{\lambda}\right) \right\} - \right. \\ \qquad \left. - 5\alpha \left\{ 1 - \frac{9}{20}\left(\frac{a\lambda'}{\lambda}\right)^2 + \frac{3}{5}\left(\frac{a^2\lambda''}{\lambda}\right) - \frac{1}{10}\left(\frac{a^2\lambda''}{\lambda}\right)\left(\frac{a\lambda'}{\lambda}\right) \right\} \right] . \end{cases}$$

The degree of covalency λ is a function of the distance between the ions and is estimated from the chemical shift of the halogen nucleus. Assuming $\lambda \propto \exp[-R/0.345]$, where R is the interionic distance in Å units, such that $\lambda'/\lambda = -1/0.345$ and $\lambda''/\lambda = (1/0.345)^2$, YOSIDA and MORIYA have calculated the coefficient $N_{\mu n}$.

The relaxation time τ can be defined by $\tau - 1/\overline{P}$, where $\overline{P}$ is the average value of P, averaged over all values of m and μ as the relaxation time in the presence of quadrupolar interaction is not unique. However, KRANENDONK's values of relaxation times in the absence of covalent and polarization effects ($\gamma = 1$) are about 4 to 6 orders of magnitude larger than those observed. In order to account for the experimental relaxation times, he assumes multiplication factor γ of the order of $10^2 \div 10^3$. The values of nuclear relaxation times in ionic crystals calculated by YOSIDA and MORIYA on the basis of covalent bonding effects alone are about 3 to 10 times larger than the observed ones. This discrepancy is traced in the simple Debye treatment of the lattice vibrations and ambiguities in the determinations of the Debye temperature θ or sound velocity v. The interaction of the covalent electron with the quadrupole moment is assumed to be unaffected by shielding and antishielding effects and the relaxation time is attributed to covalent effects alone.

KRAUS and TANTTILA ([29]) have computed the probability per unit time that the sound waves will produce a transition of a nuclear spin between its Zeeman levels. The nuclear quadrupolar interaction Hamiltonian is given by

$$\mathcal{H} = \sum_{\alpha} \frac{q_\alpha}{(r^\alpha)^5} \sum_{i=-2}^{2} C_i Q_i B_{-i}^{(\alpha)}, \tag{25}$$

where q_α is an external charge to a nucleus N, $r^{(\alpha)}$ is the distance between the nuclear center of mass and the external charge q_α when both of them are displaced by sound waves, C_i are numerical constants, Q_i are the components of the nuclear quadrupole moment given by $Q_i = \int \varrho_n(r) A_i \, d\tau_n$, $(i = 0, \pm 1, \pm 2)$. The quadrupolar term of the electrostatic potential at a position r_n from the center of mass of the nucleus, where ϱ_n is the nuclear charge density due to the charge q_α after displacement, is given by

$$V_\alpha = \frac{q_\alpha}{6(r^{(\alpha)})^5} \boldsymbol{A} : \boldsymbol{B}, \tag{26}$$

where $\boldsymbol{A}$ and $\boldsymbol{B}$ are second rank symmetric tensors.

([29]) O. KRAUS and W. H. TANTTILA: *Phys. Rev.*, **109**, 1052 (1958).

The matrix elements of the operator $\mathcal{H}$ which are of interest are given by

$$(27)\qquad \begin{cases} (I_m|\mathcal{H}|I_{m+2}) = (I_m|Q_{-2}|I_{m+2}) \sum_\alpha \dfrac{q_\alpha}{R_\alpha^5} C_2 B_{+2}^{(\alpha)}, \\ (I_m|\mathcal{H}|I_{m-2}) = (I_m|Q_{+2}|I_{m-2}) \sum_\alpha \dfrac{q_\alpha}{R_\alpha^5} C_2 B_{-2}^{(\alpha)}. \end{cases}$$

The displacement of a lattice point in the presence of unidirectional standing sound waves being propagated along the x-axis is given by

$$S = A \cos(kx - \delta) \cos \omega t,$$

where $k = 2\pi/\lambda$ with λ the wavelength of the sound waves, A the amplitude of the sound waves and δ an arbitrary phase factor. The displacement of the nucleus is taken as $S_1^{(0)} = A \cos\delta \cos\omega t$ and for the charge q_α as $S_1^{(1)} = A \cos(ka - \delta) \cos \omega t$.

Expanding $B_{+2}^{\alpha}/R_\alpha^5$ and $B_{-2}^{\alpha}/R_\alpha^5$ in terms of the relative displacement of q_α and utilizing the fact that

$$(28)\qquad \begin{cases} (I_m|Q_{+2}|I_{m-2}) = \dfrac{3eQ}{I(2I-1)}[(I-m+2)(I+m-1)(I-m+1)(I+m)]^{\frac{1}{2}}, \\ (I_m|Q_{-2}|I_{m+2}) = -\dfrac{3eQ}{I(2I-1)}[(I+m+2)(I-m-1)(I+m+1)(I-m)]^{\frac{1}{2}}, \end{cases}$$

the matrix elements are given by

$$(29)\qquad \begin{cases} (I_m|\mathcal{H}|I_{m+2}) = \dfrac{-e^2Q\gamma}{8I(2I-1)} \cdot \\ \qquad \cdot[(I+m+2)(I-m-1)(I+m+1)(I-m)]^{\frac{1}{2}} \dfrac{18}{a^3} Ak \sin\delta \cos\omega t, \\ (I_m|\mathcal{H}|I_{m-2}) = \dfrac{-e^2Q\gamma}{8I(2I-1)} \cdot \\ \qquad \cdot[(I-m+2)(I+m-1)(I-m+1)(I+m)]^{\frac{1}{2}} \dfrac{18}{a^3} Ak \sin\delta \cos\omega t. \end{cases}$$

where $q_\alpha = \gamma e$.

From the first order perturbation theory the probability that in time t the nucleus will make a transition from the initial state m to one of the states k is given by

$$(30)\qquad P_{km} = \frac{|\mathcal{H}^0|^2}{\hbar^2} g(\omega) \frac{\pi}{2} t,$$

where $\mathcal{H}^0$ is time independent and $g(\omega)$ is the normalized shape function of the nuclear resonance line. Averaging over the frequency spread $\delta\omega$ of the resonance line the average transition probability per unit time $\bar{P}$ is given by

$$\bar{P} = |\mathcal{H}^0_{km}|^2/4\hbar^2\,\delta\nu\,. \tag{31}$$

Thus the transition probability per unit time for $\Delta m = \pm 2$ transitions is given by

$$\left.\begin{aligned} P(m, m-2) &= \frac{81}{64\hbar^2\delta\nu}\,\frac{e^4Q^2\gamma^2}{I^2(2I-1)^2}\cdot \\ &\quad\cdot(I-m+2)(I+m-1)(I-m+1)\,(I+m)\,\frac{A^2k^2}{a^6}\,\sin^2\delta\,, \\ P(m, m+2) &= \frac{81}{64\hbar^2\delta\nu}\,\frac{e^4Q^2\gamma^2}{I^2(2I-1)^2}\cdot \\ &\quad\cdot(I+m+2)(I-m-1)(I+m+1)\,(I-m)\,\frac{A^2k^2}{a^6}\,\sin^2\delta\,. \end{aligned}\right\} \tag{32}$$

Kraus and Tanttila have also derived expressions for the equilibrium nuclear magnetisation in the presence of ultrasonic and thermal lattice vibrations. The time rate-of-change of the z-component, m_z, of the macroscopic magnetisation, which arises from the nuclear spin system is given by

$$\frac{\mathrm{d}m_z}{\mathrm{d}t} = \left.\frac{\mathrm{d}m_z}{\mathrm{d}t}\right|_{\Delta m=\pm1} + \left.\frac{\mathrm{d}m_z}{\mathrm{d}t}\right|_{\Delta m=\pm2} + \left.\frac{\mathrm{d}m_z}{\mathrm{d}t}\right|_{\text{ultrasonic}}, \tag{33}$$

where

$$\left.\frac{\mathrm{d}m_z}{\mathrm{d}t}\right|_{\Delta m=\pm1} = \frac{6(m_{z_0}-m_z)}{I(I+1)\,(2I+1)}\sum_{m=I}^{-I+1}\omega^m_{m-1}\,,$$

$$m_{z_0} = \frac{1}{3}\gamma\hbar I(I+1)\,n(\hbar\omega/KT_l)\,,$$

$= $ value of m_z when the spin system is in thermal equilibrium with the lattice,

$T_l =$ Temperature of the lattice.

Here ω^m_{m-1} is the probability per unit time for a thermally induced transition of a spin from the state m to the state $m-1$ and given by

$$\omega^m_{m-1} = \frac{9e^2Q^2}{4I^2\,(2I-1)^2}\,(2m-1)^2(I+m)\,(I-m+1)\,F(l)\,, \tag{34}$$

where $F(l)$ is a function of the lattice co-ordinates only.

Similarly $(\mathrm{d}m_z/\mathrm{d}t)|_{\Delta m=\pm 2}$ is given by

$$\left.\frac{\mathrm{d}m_z}{\mathrm{d}t}\right|_{\Delta m=\pm 2} = \frac{24(m_{z_0}-m_z)}{I(I+1)(2I+1)} \sum_{m=I}^{-I+2} \omega_{m-2}^{m}, \tag{35}$$

where

$$\omega_{m-2}^{m} = \frac{9e^2Q^2}{I^2(2I-1)^2}(I+m)(I-m+1)(I+m-1)(I-m+2)G(l),$$

and again $G(l)$ is a function of the co-ordinates only.

If the frequency of the sound waves is chosen in such a way that it produces $\Delta m = \pm 2$ transitions

$$\left.\frac{\mathrm{d}m_z}{\mathrm{d}t}\right|_{\text{ultrasonic}} = \frac{-24}{I(I+1)(2I+1)} m_z \sum_{m=I}^{-I+2} P_{m-2}^{m}, \tag{36}$$

where the ultrasonic probabilities are given by equations (32).

Introducing relaxation times $\mathrm{d}m_z/\mathrm{d}t$ can be expressed as

$$\frac{\mathrm{d}m_z}{\mathrm{d}t} = \frac{m_{z_0}-m_z}{T_1} - \frac{m_z}{T_u}, \tag{37}$$

where

$$\left\{\begin{aligned} &\frac{1}{T_1} = \frac{1}{T'} + \frac{1}{T''}, \\ &\frac{1}{T'} = \frac{6}{I(I+1)(2I+1)} \sum_{m=I}^{-I} \omega_{m-1}^{m}, \\ &\frac{1}{T''} = \frac{24}{I(I+1)(2I+1)} \sum_{m=I}^{-I+2} \omega_{m-2}^{m}, \\ &\frac{1}{T_u} = \frac{24}{I(I+1)(2I+1)} \sum_{m=I}^{-I+2} P_{m-2}^{m}. \end{aligned}\right. \tag{38}$$

Here T_1 is the observed spin-lattice relaxation time. Ultrasonic relaxation time can be computed from equations (38) and (32).

If the ultrasonic excitation which is applied to the specimen is constant and the time through which it is applied is longer than T_1, the magnetization is constant in a region where $\delta =$ constant. Thus putting $\mathrm{d}m_z/\mathrm{d}t = 0$, we obtain

$$m_z = \frac{m_{z_0}}{1+(T_1/T_u)} = \frac{m_{z_0}}{1+\beta \sin^2\delta}, \tag{39}$$

where

$$\beta = \frac{T_1}{\sin^2\delta} \frac{24}{I(I+1)(2I+1)} \sum_{m=I}^{-I+2} P_{m-2}^{m} . \tag{40}$$

The total magnetisation is obtained by integrating equation (39) over the volume of the crystal and is given by

$$M_z = \int m_z \, \mathrm{d}T = m_{z_0} \int \frac{A \, \mathrm{d}x}{1 + \beta \sin^2 [(2\pi/\lambda)x]} = \frac{M_{z_0}}{(1+\beta)^{\frac{1}{2}}} , \tag{41}$$

where A is the cross-sectional area of the uniform cylinder, $\delta = 2\pi x/\lambda$, λ = ultrasonic wavelength, M_{z_0} = total magnetisation of nuclei in the absence of ultrasonic excitation.

Abragam and Proctor ([30]) have also deduced the ratio (M_z/M_{z_0}) without going into the exact expressions for the contributions to the rates of change of populations from the spin-lattice relaxation and spin-spin coupling. The rates of change of population of the levels $I_z = m$ of ^{23}Na irradiated by ultrasonic waves at twice the Larmor frequency are given by

$$\left\{ \begin{aligned} \frac{\mathrm{d}n_{\frac{3}{2}}}{\mathrm{d}t} &= -W(n_{\frac{3}{2}} - n_{-\frac{1}{2}}) + F_{\frac{3}{2}}(n_m) + G_{\frac{3}{2}}(n_m) , \\ \frac{\mathrm{d}n_{\frac{1}{2}}}{\mathrm{d}t} &= -W(n_{\frac{1}{2}} - n_{-\frac{3}{2}}) + F_{\frac{1}{2}}(n_m) + G_{\frac{1}{2}}(n_m) , \\ \frac{\mathrm{d}u_{-\frac{1}{2}}}{\mathrm{d}t} &= -W(n_{-\frac{1}{2}} - n_{\frac{3}{2}}) + F_{-\frac{1}{2}}(n_m) + G_{-\frac{1}{2}}(n_m) , \\ \frac{\mathrm{d}n_{-\frac{3}{2}}}{\mathrm{d}t} &= -W(n_{-\frac{3}{2}} - n_{\frac{1}{2}}) + F_{-\frac{3}{2}}(n_m) + G_{-\frac{3}{2}}(n_m) , \end{aligned} \right. \tag{42}$$

where F_m and G_m represent the contributions to the rates of change of the populations from the spin-lattice relaxation and spin-spin coupling, respectively. Writing $\langle M_z \rangle = \frac{3}{2}(n_{\frac{3}{2}} - n_{-\frac{3}{2}}) + \frac{1}{2}(n_{\frac{1}{2}} - n_{-\frac{1}{2}})$, the contribution of spin-lattice relaxation to the rate of change of $\langle M_z \rangle$, *i.e.*,

$$\frac{3}{2} F_{\frac{3}{2}}(\,) + \ldots + \left(-\frac{3}{2}\right) F_{-\frac{3}{2}} = -\frac{M_z - M_{z_0}}{T_1} ,$$

provided a single spin-lattice relaxation is assumed, we obtain

$$\frac{\mathrm{d}\langle M_z \rangle}{\mathrm{d}t} = -\frac{\langle M_z \rangle - M_{z_0}}{T_1} - 2W\{(n_{\frac{3}{2}} - n_{-\frac{3}{2}}) + (n_{\frac{1}{2}} - n_{-\frac{1}{2}})\} , \tag{43}$$

([30]) A. Abragam and V. G. Proctor: *Phys. Rev.*, **109**, 1441 (1958).

as the contributions due to spin-spin interaction is zero because of conservation of energy. As Boltzmann distribution is maintained throughout

$$(n_{\frac{3}{2}} - n_{-\frac{3}{2}}) = 3(n_{\frac{1}{2}} - n_{-\frac{1}{2}}) \quad \text{and consequently} \quad \langle M_z \rangle = \tfrac{10}{6}(n_{\frac{3}{2}} - n_{-\frac{3}{2}})$$

or

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle M_z \rangle = -\frac{\langle M_z \rangle - M_{z_0}}{T_1} - \frac{8W}{5}\langle M_z \rangle . \tag{44}$$

The limiting value of $\langle M_z \rangle$ is given by putting $\langle \mathrm{d}M_z/\mathrm{d}t \rangle = 0$

$$\langle M_z \rangle / M_{z_0} = (1 + 8WT_1/5)^{-1} .$$

Since the signal amplitude ξ is proportional to $\langle M_z \rangle$, we have

$$\xi/\xi_0 = \langle M_z \rangle / M_{z_0} = \left(1 + \frac{8WT_1}{5}\right)^{-1} . \tag{45}$$

3. – Experimental technique.

The technique is essentially based upon the pulsed nuclear induction method of measuring nuclear magnetization as described by HAHN ([31]). The thing to be remembered is that if a short pulse of radio frequency magnetic field is applied at the Larmor frequency, the transient nuclear induction signal is proportional in magnitude to the static magnetization of the system, which is a function of the population of the various states. Thus the amplitude of the transient nuclear induction signal A_1 following the pulse of radio frequency flux is proportional to the population difference between the quadrupole states at thermal equilibrium. Once the equilibrium is disturbed the instantaneous population difference Δ will try to reach its equilibrium value Δ_e at a rate which is determined by the thermal relaxation time T_1 according to the relation

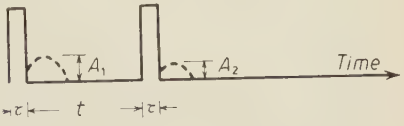
$$\frac{\Delta(t)}{\Delta_e} = (1 - \exp[-t/T_1]) , \tag{46}$$

provided the time interval between the two successive radio frequency pulses is greater than the thermal relaxation time T_1 and $\Delta(0) = 0$ which is the case when the pulse length τ is adjusted to give maximum transient in-

([31]) E. L. HAHN: *Phys. Rev.*, **77**, 297 (1949).

duction signal. If A_2 is the amplitude of the transient nuclear induction signal after a second pulse of radio frequency flux of the same duration is applied after a time t (see Fig. 1), then

$$(47) \qquad \frac{A_2}{A_1} = (1 - \exp[-t/T_1]) .$$



By measuring the ratio (A_2/A_1) as a function of the time interval t and plotting $\log_e (1 - A_2/A_1)$ *vs.* t, the thermal relaxation T_1 is known from the slope of the straight line.

Fig. 1.

Next the ultrasonic excitation is applied at twice the Larmor frequency and the magnetization is again measured. The typical time sequence of events is shown in Fig. 2. If ξ_0 is the amplitude of the nuclear induction signal in the absence of ultrasonic excitation and ξ, in the presence of ultrasonic excitation, the ratio can be expressed as

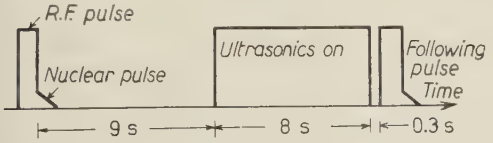


Fig. 2.

$$(48) \qquad \left(\frac{\xi_0}{\xi}\right)^2 = 1 + k_1 V^2 ,$$

where V is the peak voltage applied to the transducer. If we put $M_{z_0} = \xi_0$, $M_z = \xi$ and $\mathrm{d}M_z/\mathrm{d}t = 0$ in equation (41), we obtain

$$(49) \qquad k_1 V^2 = \beta_1 = \frac{243}{8a^6\hbar^2\delta\nu} \frac{e^4 Q^2 \gamma^2 A^2 k^2}{I^3(I+1)(2I+1)(2I-1)^2} T_1 \cdot \\ \cdot (I-m+2)(I+m-1)(I-m+1)(I+m) ,$$

where $\delta\nu$ is the ultrasonic line width, A is the amplitude of ultrasonic vibrations, $k = 2\pi a/\lambda$, λ is the ultrasonic wavelength, a is the lattice constant, e is the electronic charge and the nearest neighbours of a given nucleus are replaced by point charges of magnitude γe.

Thus from the experimentally measured values of k_1 (which is known from the slope of the curve $(\xi_0/\xi)^2$ *vs.* V^2), V, $\delta\nu$ (which is known from the plot of (ξ/ξ_0) *vs.* ultrasonic frequency at constant power around $2\nu_0$, ν_0 being the nuclear magnetic resonance frequency), k and T_1, one can obtain the information about the parameter γ provided the amplitude of ultrasonic vibrations in the crystal is known. Ultrasonic energy density in the crystal is calculated from the power transmitted to the crystal and the phonon relaxation times of the ultrasonic phonons from the relation

$$(50) \qquad V' \mathcal{E} = P_s T_\varrho ,$$

where V' is the volume of the specimen, $\mathcal{E}$ is the ultrasonic energy density, P_s is the power delivered to the specimen, T_g the sound attenuation coefficient or phonon relaxation time. The power delivered to the sample is known in terms of the voltage across the quartz and is given by

$$P_s = \frac{R_e^{\text{eff}}}{(R_u^{\text{eff}} + R_e^{\text{eff}})^2} V^2 , \tag{51}$$

where R_e^{eff} is the effective load resistance and R_u^{eff} is the effective internal resistance of the quartz and its circuits. The sound attenuation coefficients are known from measurements of HUNTINGTON ([32]) and GALT ([33]). PROCTOR and ROBINSON ([34]) have also measured it by studying the attenuation of the nuclear induction signal as a function of the frequency of ultrasonic excitation both when one face of the specimen is or is not in contact with castor oil. In the former case since some of the acoustic energy escapes into oil, the attenuation of the nuclear induction signal is less. From a knowledge of acoustic impedances of NaCl and castor oil and from the loss of attenuation of the nuclear induction signal they calculated the value of 10 MHz phonon relaxation time which is in good agreement with the values obtained by HUNTINGTON and GALT.

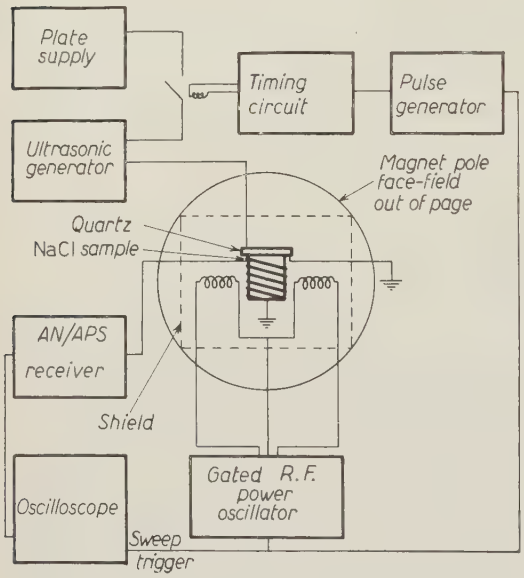


Fig. 3.

The block diagram of the apparatus which has been used by PROCTOR and his associates is given in Fig. 3. It consists of two main groups, one which is used in producing ultrasonic excitation and the other which measures the nuclear magnetization. A nuclear induction head of the type described by WEAVER ([35]) is used. The transmitter and the receiving coils are at right angles to each other and a static magnetic field is applied perpendicular to

([32]) H. B. HUNTINGTON: *Phys. Rev.*, **72**, 311 (1947).
([33]) J. K. GALT: *Phys. Rev.*, **73**, 1460 (1948).
([34]) W. G. PROCTOR and W. A. ROBINSON: *Phys. Rev.*, **104**, 1344 (1956).
([35]) H. E. WEAVER: *Phys. Rev.*, **89**, 923 (1953).

the axes of both the coils. The quartz crystal is glued to one end of an NaCl rod of about $\frac{1}{2}$ in. diameter and one in. length which serves the purpose of an extension the of ultrasonic path length for convenience. The other end of this rod is cemented to the sample crystal of the same diameter. The sample crystal is well insulated from the ultrasonic circuits. PROCTOR and ROBINSON used a steady magnetic field of 4220 oersteds and a short pulse of 50 μs of radio frequency magnetic field at the Larmor frequency 4.75 MHz for studying the transient radio frequency nuclear signal whose magnitude is proportional to the nuclear magnetization just before the pulse. Ultrasonic excitation was applied nine seconds after the preceding nuclear magnetization measurement, at twice the Larmor frequency for 8 seconds and after a delay of 0.03 second the nuclear magnetization was again measured (see Fig. 2).

4. – Results and discussions.

The data on ultrasonic investigation of nuclear spin-lattice relaxation time are very limited. The only crystals in which ultrasonic excitation has been used to study its effects on nuclear spin-lattice relaxation are NaCl, NaI, $NaClO_3$ and InSb. $NaClO_3$ was the first crystal to be studied by PROCTOR and TANTTILA. They observed that the population difference of the two pure quadrupole levels decreased when the sample was subjected to ultrasonic energy at the transitions corresponding to $\Delta m = \pm 1$ and $\Delta m = \pm 2$. However, it was realised that the magnetic fields due to currents in the ultrasonic system might affect the population difference. Consequently PROCTOR and ROBINSON used ultrasonic excitation corresponding to $\Delta m = \pm 2$ transitions in the case of NaCl and the specimen too was placed in a static magnetic field.

Assuming that the quadrupolar relaxation is the most important mechanism in NaCl, PROCTOR and ROBINSON calculated γ on the basis of Kranendonk's six point model. The value of γ comes out to be equal to 3 which is somewhat lower than the value of 50 estimated by KRANENDONK. However, as he points out, his value is a measure of the second derivative of the electric field gradient whereas the value of γ which is calculated from ultrasonic measurements is a measure of the first derivative. The discrepancy can also be attributed to the lack of reliable measurements of ultrasonic attenuation as reported in the literature and in the ambiguity in equating phonon relaxation times with ultrasonic attenuation. In view of the fact that the knowledge of the amplitude of ultrasonic vibrations in the crystal helps a great deal in determining the strength of the coupling, it should be known more accurately than what is possible now. However, if one does not care for the absolute value of the quadrupolar coupling the problem of determining the correct

539

value of ultrasonic energy or amplitude in the crystal can be avoided by comparing the net magnetizations of two sets of nuclei in equilibrium with thermal and lattice vibrations caused by ultrasonic waves in the direction of a constant external magnetic field H_0. This was done in the case of NaI crystal by JENNINGS, TANTTILA and KRAUS [36]. They determined experimentally the equilibrium magnetizations in the direction of the static external magnetic field of sodium and iodine nuclei as a function of the ultrasonic power delivered to the sodium iodide crystal which in turn is proportional to the square of the peak voltage applied to the quartz crystal. By plotting $(\xi_0/\xi)^2$ *vs.* V^2 for both sodium and iodine nuclei, the slopes k_{Na} and k_{I} are known. However, the ratio of the two slopes is given by

$$\frac{k_{\mathrm{I}}}{k_{\mathrm{Na}}} = \frac{T_{1\mathrm{I}}}{T_{1\mathrm{Na}}} \frac{Q_{\mathrm{I}}^2 \gamma_{\mathrm{I}}^2}{\delta\nu_{\mathrm{I}}} \cdot \frac{\delta\nu_{\mathrm{Na}}}{Q_{\mathrm{Na}}^2 \gamma_{\mathrm{Na}}^2} (0.240) , \tag{52}$$

provided we assume that the amplitude of ultrasonic vibrations A is proportional to the applied voltage ($A = CV$, where C is a constant). JENNINGS, TANTTILA and KRAUS obtain the following values experimentally: $T_{1\mathrm{I}} = 0.0065$ s, $T_{1\mathrm{Na}} = 5$ s, $(k_{\mathrm{I}}/k_{\mathrm{Na}}) = 2.2$, $\delta\nu_{\mathrm{Na}} = 4.65$ kHz/s, $\delta\nu_{\mathrm{I}} = 4.49$ kHz/s. Substituting these values together with values for quadrupole moments for sodium and iodine nuclei, they obtain

$$(\gamma_{\mathrm{I}}/\gamma_{\mathrm{Na}}) = 10.9 ,$$

which is in qualitative agreement with the theoretical calculations of STERNHEIMER [37] and DAS and BERSOHN [38]. However, k is to be compared with $1+\gamma_\infty$ where γ_∞ is the polarization calculated by STERNHEIMER. According to Sternheimer's calculations, γ_∞ for sodium ions is 4.2. Though he does not give any calculation for iodide ion, yet it can be guessed from the corresponding value in the cesium ion, in which case he obtains $\gamma_\infty = 143$. Supposing that the γ_∞ for the iodide ions is not less than 143, $(\gamma_{\mathrm{I}}/\gamma_{\mathrm{Na}}) = 28$ according to Sternheimer's calculations. The discrepancy between the two results can, however, be traced in the simplified theoretical model which has been used by JENNINGS, TANTTILA and KRAUS. Moreover, imperfections in the crystal which are known very well to cause broadening effects on the satellite lines of iodine nuclei should also be taken into consideration. Imperfections in the crystal give rise to random but permanent electric field gradients which interact with nuclear quadrupole moments. This might contribute in the right direction towards the increase in the ratio of $(\gamma_{\mathrm{I}}/\gamma_{\mathrm{Na}})$. However, in the treat-

[36] D. A. JENNINGS, W. H. TANTTILA and O. KRAUS: *Phys. Rev.*, **109**, 1059 (1958).
[37] R. M. STERNHEIMER: *Phys. Rev.*, **105**, 158 (1957).
[38] T. P. DAS and R. BERSOHN: *Phys. Rev.*, **102**, 733 (1956).

ment given by KRAUS and TANTTILA, the spin system has been assumed to have a uniform temperature, which may not be valid in view of the fact that ultrasonic excitation causes hot spots in the system and the spin conduction cannot be neglected.

The above calculations are based on the concept of a spin temperature as distinct from lattice temperature whose validity has been discussed in great detail by ABRAGAM and PROCTOR ([30]). The existence of spin temperature as a valid assumption leads to certain predictions which can be tested experimentally. For nuclei $I > \frac{1}{2}$ the equidistance of the levels $I_z = m$ of a spin in a magnetic field H_0 makes it possible to define the spin temperature according to the relation

$$P_m \sim \exp\{-\gamma\hbar H_0 m/kT\},$$

where the populations P_m of these levels form a Boltzmann distribution. The Boltzmann distribution of the levels is maintained and even if it is not so initially at $t = 0$, spin-spin interactions bring it to that form. The test of such an effect is provided by ultrasonic saturation experiments. Equation (45), which is derived on the assumption of Boltzmann distribution shows that M tends to zero as the ultrasonic transition probability becomes sufficiently strong. ABRAGAM and PROCTOR have been able to increase the density of ultrasonic waves in the crystal to such an extent that the magnetic moment of ^{23}Na completely disappears. Ultrasonic experiments illustrate the role of the spin-spin interaction in establishing and maintaining a Boltzmann distribution between the populations of the spin system.

PROCTOR and ROBINSON have also attempted to seek information regarding the number of imperfections in the crystal from ultrasonic line width. ROBINSON ([39]) has shown that the average field gradient caused by dislocations is approximately given by $q \sim 6e\gamma\zeta^{\frac{1}{2}}/a^2$, where ζ is the number of dislocations per square centimeter. Equating the line width to twice the pure quadrupole interaction energy $eqQ/2$, the line width is given by

$$\Delta\nu \sim \frac{6e^2\gamma Q\zeta^{\frac{1}{2}}}{a^2 h}.$$

PROCTOR and ROBINSON obtain the ultrasonic line width in NaCl equal to 4 kHz. Using this value ζ comes out to be equal to $\sim 10^{10}$ dislocations/cm^2, which is in fair agreement with the values obtained by other methods.

MENES and BOLEF ([40]) studied the absorption of acoustic energy at the

([39]) W. A. ROBINSON: *Thesis* (University of Washington, 1956), unpublished.
([40]) M. MENES and D. I. BOLEF: *Phys. Rev.*, **109**, 218 (1958).

resonant frequency of ^{115}In nuclei in a single crystal of InSb at both the normal nuclear magnetic resonance frequency ($\Delta m = \pm 1$) and at the double frequency ($\Delta m = \pm 2$). The nuclear resonance acoustic absorption was studied by the change in the mechanical Q of the sample, which changed the electrical impedance of the transducer which controlled the Q of the tank circuit of a modified Pound-Watkins type of spectrometer.

So far the ultrasonic excitation has been applied in only one fixed direction, say the 100 direction, but if it is applied in different directions (*), the picture which we may get with regard to anisotropy of γ in the crystal may yield interesting information on electron distributions and their relative displacement under ultrasonic excitation. It will be further interesting if ultrasonic excitation is applied to solids where spin-lattice relaxation is due to magnetic dipole interaction. So far as the improvements in the technique are concerned, some better ways are yet to be devised which may give reliable values of amplitudes of ultrasonic waves in the crystal or displacements of lattice points under ultrasonic excitation.

* * *

Part of this work was carried out while the author was at Western Reserve University, Cleveland, Ohio, U.S.A. The author gratefully acknowledges a grant from Office of Naval Research which made his stay in Cleveland possible. The author is also grateful to Professor Ernest Yeager for his helpful comments and suggestions.

(*) *Note added in proof.* – Recently Taylor and Bloembergen (*Phys. Rev.*, **113**, 431 (1959), have studied nuclear spin saturation by Ultrasonics in NaCl crystal. They found no angular dependence in the saturation curves. They have been successful in measuring the components of the fourth order tensor S, which connects the electric field gradient tensor at the nucleus with the strain deformation tensor ε

$$\nabla_i \nabla_j V = \sum_{k,l} S_{ijkl} \varepsilon_{kl} , \qquad (i, j, k, l = x, y, z) .$$

Absence of angular dependence implies that $S_{44}/S_{11} = \frac{3}{4}$, which differs from Cauchy relation $S_{44}/S_{11} = -\frac{1}{2}$. Also the experimental ratio $(S_{11})_{Ce}/(S_{11})_{Na} = 1.8$, whereas the ionic point-charge model coupled with an isotropic Sternheimer antishielding factor gives a factor 10. The results show inadequacy of ionic point-charge model and indicate the possibility of a considerable amount of covalent character and configurational interaction on the Na^+ ion.

SUPPLEMENTO AL VOLUME XII, SERIE X
DEL NUOVO CIMENTO

N. 1, 1959
2° Trimestre

Die Zeit in der Quantenmechanik.

F. Engelmann und E. Fick

Institut für Theoretische Physik der Technischen Hochschule - München

(ricevuto il 22 Dicembre 1958)

Inhalt. — **1.** Einleitung. – **2.** Bildunabhängige Bewegungsgleichungen im Hilbert-Raum. – **3.** Spezielle Bezugssysteme. – **4.** Der Zeitoperator. – **5.** Diskussion.

1. – Einleitung.

Die Beschreibung der zeitlichen Veränderung des Zustands eines quantenmechanischen Systems erfolgt in der Schrödinger-Darstellung durch die Bewegung des Zustandsvektors in einem Hilbert-Raum $\mathfrak{H}$, während in der Heisenberg-Darstellung der Zustandsvektor zeitlich konstant gehalten wird und die Operatoren in $\mathfrak{H}$ gerade eine solche Bewegung erfahren, daß in beiden Darstellungen die relative Bewegung des Zustandsvektors gegenüber den Operatoren dieselbe ist. Durch eine unitäre Transformation ist es bekanntlich möglich, von einem Bild in das andere zu gelangen.

In der vorliegenden Arbeit wollen wir darauf hinweisen, daß eine bildunabhängige Formulierung der Bewegungsgleichung des Zustandsvektors möglich ist, die man unmittelbar aus der statistischen Deutung der Quantenmechanik erhält, wenn man die allgemeine Quantisierungsvorschrift, d.h. die Übersetzung von Observablen in Operatoren hinzunimmt. Die zeitlich aufeinanderfolgenden Zustände werden in der Quantenmechanik bekanntlich durch einen Zeitparameter t beschrieben, der eine gewöhnliche Zahl (c-Zahl) und kein Operator ist. Die bildunabhängige Formulierung der Bewegungsgleichungen führt uns darauf, daß neben dieser c-Zahl-Zeit t noch ein Zeitoperator $\mathfrak{T}$

für die Observable « Zeit » (1) einzuführen ist, dessen Erwartungswert t ist. $\mathcal{T}$ ist explizit unabhängig von t, falls der Hamilton-Operator explizit von t nicht abhängt. Durch diesen Operator wird es möglich, die Unbestimmtheitsrelation zwischen Energie und Zeit direkt zu beweisen, was in der üblichen Formulierung der Quantenmechanik, in der die Zeit nur als c-Zahl vorkommt, auf Schwierigkeiten führt.

2. – Bildunabhängige Bewegungsgleichungen im Hilbert-Raum.

Um die Bedeutung der Zeit in der Quantenmechanik verstehen zu können, ist eine genaue Analyse der Quantisierung erforderlich. Zur quantenmechanischen Behandlung der physikalischen Erscheinungen gelangt man, wenn man von den folgenden Forderungen ausgeht:

1.) Es besteht eine ein-eindeutige Zuordnung zwischen den Observablen F und den hermitischen Operatoren $\mathcal{F}$ eines Hilbert-Raums $\mathfrak{H}$.

2.) Klassische Beziehungen zwischen den Observablen gelten für die zugeordneten Operatoren, wenn man im Rahmen der Hamilton-Jacobischen Theorie die Poisson-Klammern $\{FG\}$ durch die Kommutatoren $(i/h)[\mathcal{F}\mathcal{G}]_-$ ersetzt und Produkte symmetrisiert.

3.) Es soll eine statistische Interpretation möglich sein; d.h. es existiert ein statistischer Operator, der den Zusammenhang zwischen den Operatoren und den Erwartungswerten der Observablen vermittelt (2).

Die Übersetzung der Hamilton-Jacobischen Theorie liefert gemäß Forderung 1.) und 2.) für jede Observable F eine Operatorfunktion

$$\mathcal{F} = \mathcal{F}(q_i, p_i, t) , \tag{1}$$

wobei q_i und p_i die Orts- und Impulsoperatoren darstellen, während die Zeit t, sofern sie explizit auftritt, lediglich als eine gewöhnliche Zahl (c-Zahl-Parameter) eingeht. Es erweckt dies den Anschein, als ob der Observablen « Zeit » im Gegensatz zur Forderung 1.) kein Operator zugeordnet wird. Wir werden jedoch im Abschnitt **4** zeigen, daß ein Zeitoperator existiert, obwohl in (1) die Zeit nur als c-Zahl auftritt.

(1) Wir beschränken uns in der vorliegenden Arbeit auf die nichtrelativistische Quantentheorie. In einer relativistischen Formulierung übertragen sich die Überlegungen auf die Eigenzeit.

(2) J. VON NEUMANN: *Mathematische Grundlagen der Quantenmechanik* (Berlin, 1932); G. LUDWIG: *Die Grundlagen der Quantenmechanik* (Berlin-Göttingen-Heidelberg, 1954).

Auf Grund der Forderung 2.) gilt, wenn H die Hamiltonfunktion ist, die Korrespondenz

$$\dot{F} = \{H, F\} + \left(\frac{\partial F}{\partial t}\right)_{p,q} \rightarrow \boxed{\overset{0}{\mathcal{F}} = \frac{i}{\hbar}[\mathcal{H}\mathcal{F}] + \left(\frac{\partial \mathcal{F}}{\partial t}\right)_{p,q}}. \tag{2}$$

Dabei bezeichnen wir entsprechend der Zuordnung

$$D \equiv \frac{\mathrm{d}F}{\mathrm{d}t} \equiv \dot{F} \quad \rightarrow \quad \mathcal{D} \equiv \frac{\delta \mathcal{F}}{\delta t} \equiv \overset{0}{\mathcal{F}} \tag{3}$$

mit $\overset{0}{\mathcal{F}} = \mathcal{D}$ den Operator, der nach 1.) der Observablen $\dot{F} \equiv D$ entspricht (z.B. $m\overset{0}{q} = p$). $\overset{0}{\mathcal{F}}$ beschreibt also keine zeitliche Veränderung (« Bewegung ») von $\mathcal{F}$ in $\mathfrak{H}$ im Gegensatz zu $(\partial \mathcal{F}/\partial t)_{p,q}$, das die Differentiation von $\mathcal{F}$ nach seiner expliziten Zeitabhängigkeit bei festgehaltenen Operatoren p und q bedeutet ([3]). Eine Bewegung der Operatoren in $\mathfrak{H}$ kommt also lediglich durch deren explizite Zeitabhängigkeit zustande. Die Bewegung von $\overset{0}{\mathcal{F}}$ ist nach Gl. (2) durch diejenige von $\mathcal{F}$ bestimmt.

Die Gültigkeit der Gl. (2) ist an die Existenz eines Hamilton-Operators $\mathcal{H}$ geknüpft (« reversible Bewegungen »). Sie gilt deshalb bekanntlich nicht für ein Teilsystem, das Wechselwirkungen mit weiteren quantenmechanischen Systemen hat, und auch nicht für Messungen.

Ein reiner Zustand eines quantenmechanischen Systems wird durch einen Zustandsvektor Ψ in $\mathfrak{H}$ beschrieben, dem ein Projektionsoperator $\mathcal{P}_\Psi$ zugeordnet ist, der vermöge $\mathcal{P}_\Psi f \equiv (\Psi, f)\Psi$ auf Ψ projiziert und der den statistischen Operator des reinen Zustandes nach Forderung 3.) darstellt. Dem Operator $\overset{0}{\mathcal{P}}_\Psi$ entspricht als Observable die Frage, ob sich die *Zuordnung* vom statistischen Operator $\mathcal{P}_\Psi$ (bzw. vom Zustandsvektor Ψ) zum Zustand des Systems zeitlich ändert. Dies ist definitionsgemäß nicht der Fall, so daß

$$\boxed{\overset{0}{\mathcal{P}}_\Psi = \mathcal{O}} \tag{4}$$

gilt, weil unmögliche Eigenschaften dem Null-Operator zugeordnet sind. Aus denselben Gründen gilt auch für den allgemeinen statistischen Operator $\mathcal{W}$ einer gemischten Gesamtheit

$$\boxed{\cdot\, \mathcal{W} = \mathcal{O}}, \tag{5}$$

([3]) Vermöge der rechten Seite von Gl. (2) lassen sich für $\overset{0}{\mathcal{F}} \equiv \delta \mathcal{F}/\delta t$ Rechenregeln aufstellen, die den normalen Differentiationsgetzen entsprechen.

denn die Zuordnung des statistischen Operators zu dem Zustand der durch ihn beschriebenen Gesamtheit bleibt auch allgemein dauernd bestehen. Für reversible Bewegungen erhält man nun als Spezialfall der Gl. (2) die Bewegungsgleichung des statistischen Operators in $\mathfrak{H}$ als

$$\frac{i}{\hbar}[\mathcal{H}\mathcal{W}]_- + \left(\frac{\partial \mathcal{W}}{\partial t}\right)_{p,q} = \mathcal{O} \tag{6a}$$

für ein Gemisch [4] und speziell

$$\frac{i}{\hbar}[\mathcal{H}\mathcal{P}_\Psi]_- + \left(\frac{\partial \mathcal{P}_\Psi}{\partial t}\right)_{p,q} = \mathcal{O} \tag{6b}$$

für einen reinen Zustand.

Wendet man die Gl. (6b) auf irgendeinen Vektor aus $\mathfrak{H}$ an, so findet man die Bewegungsgleichung des Zustandsvektors

$$\boxed{\left(\frac{\partial \Psi}{\partial t}\right)_{p,q} = -\frac{i}{\hbar}\mathcal{H}\Psi} \,. \tag{7}$$

Es ergibt sich damit folgender Sachverhalt: Alle bisherigen Überlegungen haben sich ohne ein bestimmtes « Bild » formulieren lassen, d.h. die angeschriebenen Gleichungen sind unabhängig davon, welches Bezugssystem im Hilbert-Raum man zugrunde legt; sie gelten also *alle* sowohl etwa im Heisenberg- als auch im Schrödinger-Bild (Abschnitt 3). Als einzige Zeitabhängigkeit tritt bei allen Operatoren nur die explizite, d.h. jene bei festem p und q, auf. Für die statistischen Operatoren ist diese explizite Zeitabhängigkeit wesentlich. Gerade durch sie wird die zeitliche Änderung des Zustandes einer Gesamtheit erfaßt.

In Anwendung der obigen Gleichungen läßt sich etwa das Ehrenfestsche Theorem bildunabhängig als direkte Folge der Forderungen 1.) bis 3.) erkennen. Der Erwartungswert $\overline{\mathcal{F}}$ einer Observablen F läßt sich bekanntlich mittels des statistischen Operators schreiben:

$$\overline{\mathcal{F}} = \text{Spur}\,(\mathcal{W}\mathcal{F}) = f(t)\,. \tag{8}$$

Diese Größe ist eine vom Bezugssystem in $\mathfrak{H}$ unabhängige reine Zeitfunktion $f(t)$. Für die zeitliche Ableitung dieser Funktion ergibt sich (unabhängig von

---

[4] Diese Bewegungsgleichung der Dichtematrix entspricht bekanntlich dem Liouvilleschen Satz der klassischen Mechanik.

einer etwaigen Bewegung des Bezugssystems)

$$\dot{\overline{\mathcal{F}}} = \text{Spur}\left\{\left(\frac{\partial \mathcal{W}}{\partial t}\right)_{p,q}\cdot\mathcal{F} + \mathcal{W}\cdot\left(\frac{\partial \mathcal{F}}{\partial t}\right)_{p,q}\right\} =$$

$$= \text{Spur}\left\{-\frac{i}{\hbar}[\mathcal{H}\mathcal{W}]_-\cdot\mathcal{F} - \frac{i}{\hbar}\mathcal{W}[\mathcal{H}\mathcal{F}]_- + \mathcal{W}\overset{0}{\mathcal{F}}\right\} =$$

$$= \text{Spur}\,(\mathcal{W}\overset{0}{\mathcal{F}})\,,$$

also

$$\boxed{\dot{\overline{\mathcal{F}}} = \overline{\overset{0}{\mathcal{F}}}} \quad \text{(Ehrenfest)} \ (^5)\,. \tag{9}$$

3. – Spezielle Bezugssysteme.

Die bisher angegebenen bezugssystem-unabhängigen Gleichungen sind dadurch charakterisiert, daß in ihnen keine totalen Ableitungen $\mathrm{d}/\mathrm{d}t$ von Operatoren und Zustandsvektoren vorkommen. Diese Ableitung $\mathrm{d}/\mathrm{d}t$ einer Größe in $\mathfrak{H}$ muß sinngemäß die vollständige Änderung dieser Größe bezogen auf ein speziell gewähltes Bezugssystem in $\mathfrak{H}$ beschreiben. Erst bei Angabe eines speziellen Bezugssystems in $\mathfrak{H}$ ist also die Bedeutung von $\mathrm{d}/\mathrm{d}t$ definiert und umgekehrt läßt sich durch Angabe der Bedeutung von $\mathrm{d}/\mathrm{d}t$ ein Bezugssystem definieren.

Zum Schrödinger-Bild gelangt man durch die Forderung, daß man für die totale Zeitableitung eines Operators verlangt

$$\frac{\mathrm{d_S}\mathcal{F}}{\mathrm{d}t} = \left(\frac{\partial \mathcal{F}}{\partial t}\right)_{p,q}. \tag{10}$$

Man legt also ein Bezugssystem zugrunde, gegenüber dem die Operatoren p und q ruhen:

$$\frac{\mathrm{d_S}p}{\mathrm{d}t} = \frac{\mathrm{d_S}q}{\mathrm{d}t} = \mathcal{O}\,. \tag{11}$$

Im Heisenberg-Bild definiert man die totale Zeitableitung durch

$$\frac{\mathrm{d_H}\mathcal{F}}{\mathrm{d}t} = \overset{0}{\mathcal{F}}\,, \tag{12}$$

so daß in diesem Bezugssystem der statistische Operator ruht

$$\frac{\mathrm{d_H}\mathcal{W}}{\mathrm{d}t} = \mathcal{O}\,. \tag{13}$$

(5) Der Erwartungswert von $\mathcal{W}$ ist also zeitlich konstant (allgemeines « Integral der Bewegung »).

Die bildunabhängigen Bewegungsgleichungen liefern damit die bekannten Gleichungen im

	Schrödinger-Bild	Heisenberg-Bild
Def.:	$\frac{d_S \mathcal{F}}{dt} \equiv \left(\frac{\partial \mathcal{F}}{\partial t}\right)_{p,q}$	$\frac{d_H \mathcal{F}}{dt} \equiv \frac{i}{\hbar}[\mathcal{H}\mathcal{F}]_- + \left(\frac{\partial \mathcal{F}}{\partial t}\right)_{p,q}$
(2)	$\overset{0}{\mathcal{F}} = \frac{i}{\hbar}[\mathcal{H}\mathcal{F}]_- + \frac{d_S \mathcal{F}}{dt}$,	$\overset{0}{\mathcal{F}} = \frac{d_H \mathcal{F}}{dt}$,
(6)	$\frac{d_S \mathcal{W}}{dt} = \left(\frac{\partial \mathcal{W}}{\partial t}\right)_{p,q} = -\frac{i}{\hbar}[\mathcal{H}\mathcal{W}]_-$,	$\frac{d_H \mathcal{W}}{dt} = \frac{i}{\hbar}[\mathcal{H}\mathcal{W}]_- + \left(\frac{\partial \mathcal{W}}{\partial t}\right)_{p,q} = 0$,
(7)	$\frac{d_S \Psi}{dt} = \left(\frac{\partial \Psi}{\partial t}\right)_{p,q} = -\frac{i}{\hbar}\mathcal{H}\Psi$	$\frac{d_H \Psi}{dt} = \frac{i}{\hbar}\mathcal{H}\Psi + \left(\frac{\partial \Psi}{\partial t}\right)_{p,q} = 0$.

Schließlich läßt sich das Wechselwirkungsbild, das von einer Aufspaltung von $\mathcal{H} = \mathcal{H}^0 + \mathcal{H}^1$ ausgeht, natürlich ebenfalls aus der bildunabhängigen Darstellung gewinnen:

	Wechselwirkungsbild
Def.:	$\frac{d_W \mathcal{F}}{dt} \equiv \frac{i}{\hbar}[\mathcal{H}^0\mathcal{F}]_- + \left(\frac{\partial \mathcal{F}}{\partial t}\right)_{p,q}$
(2)	$\overset{0}{\mathcal{F}} = \frac{i}{\hbar}[\mathcal{H}^1\mathcal{F}]_- + \frac{d_W \mathcal{F}}{dt}$,
(6)	$\frac{d_W \mathcal{W}}{dt} = \frac{i}{\hbar}[\mathcal{H}^0\mathcal{W}]_- + \left(\frac{\partial \mathcal{W}}{\partial t}\right)_{p,q} = -\frac{i}{\hbar}[\mathcal{H}^1\mathcal{W}]_-$,
(7)	$\frac{d_W \Psi}{dt} = \frac{i}{\hbar}\mathcal{H}^0\Psi + \left(\frac{\partial \Psi}{\partial t}\right)_{p,q} = -\frac{i}{\hbar}\mathcal{H}^1\Psi$.

4. – Der Zeitoperator.

Zur Beschreibung der Bewegung von Operatoren und Zustandsvektoren in $\mathfrak{H}$ wird die Zeit als c-Zahl-Parameter eingeführt. Darüber hinaus muß jedoch der Observablen « Zeit » des quantenmechanischen Systems nach Forderung 1.) auch ein hermitischer Operator zugeordnet sein. Wir wollen zeigen, wie man einen solchen Operator $\mathcal{T}$ einführen kann und welcher Zusammenhang mit der c-Zahl-Zeit t besteht.

Zunächst wollen wir die vereinfachende Annahme machen, daß der Hamilton-Operator nicht explizit von t abhängt. Aus der klassischen Mechanik ist bekannt, daß man vermöge der Lösung der klassischen Bewegungsgleichungen die Zeit als Funktion von p und q in der Form

$$t - t_0 = T(p, q) - T(p_0, q_0) \tag{14}$$

darstellen kann, wobei

$$\frac{\mathrm{d}T}{\mathrm{d}t} = \{H, T\} = 1 \tag{15}$$

ist. Die Operatorfunktion, die nach den Forderungen 1.) und 2.) der klassischen Funktion T zugeordnet ist, stellt den Operator der Observablen « Zeit » dar

$$\mathcal{T} = \mathcal{T}(\boldsymbol{p}, \boldsymbol{q}), \tag{16}$$

für welchen gilt

$$\overset{0}{\mathcal{T}} = \frac{i}{\hbar}[\mathcal{H}\mathcal{T}]_- = \boldsymbol{1}. \quad (6) \tag{17}$$

Der Zeitoperator $\mathcal{T}$ ist explizit von der c-Zahl-Zeit t unabhängig, wenn dies auch für $\mathcal{H}$ gilt. Die Bedeutung der « Ring-Ableitung » kommt durch Gl. (17) besonders deutlich zum Ausdruck.

Aus Gl. (17) folgt unmittelbar, daß $\mathcal{H}$ und $\mathcal{T}$ zueinander kanonisch konjugiert sind, und somit die Unbestimmtheitsrelation zwischen Energie und Zeit

$$\Delta E \cdot \Delta T \geqslant \frac{\hbar}{2}. \tag{18}$$

(6) Das heißt im Schrödinger-Bild $\mathrm{d}_{\mathrm{S}}\mathcal{T}/\mathrm{d}t = 0$ und im Heisenberg-Bild $\mathrm{d}_{\mathrm{H}}\mathcal{T}/\mathrm{d}t = 1$.

Wenden wir uns jetzt dem allgemeinen Fall zu, daß die Hamilton-Funktion H auch explizit von t abhängt. Die Funktion T, die nunmehr auch explizit von t abhängig sein muß, genügt der Gleichung

$$\frac{\mathrm{d}T}{\mathrm{d}t} = \{H, T\} + \left(\frac{\partial T}{\partial t}\right)_{p,q} = 1 . \tag{19}$$

Die Observable «Zeit» wird durch einen Operator

$$\mathcal{T} = \mathcal{T}(p, q, t) \tag{20}$$

dargestellt, für den jetzt gilt

$$\boxed{\overset{0}{\mathcal{T}} = \frac{i}{\hbar}[\mathcal{H}\mathcal{T}]_- + \left(\frac{\partial \mathcal{T}}{\partial t}\right)_{p,q} = I} . \tag{21}$$

Man sieht, daß $\mathcal{H}$ und $\mathcal{T}$ i.a. nicht mehr zueinander kanonisch konjugiert sind.

Nach dem Ehrenfestschen Theorem [Gl. (9)]

$$\dot{\overline{\mathcal{T}}} = \overline{\overset{0}{\mathcal{T}}} = 1 \tag{22}$$

erhält man

$$\boxed{\overline{\mathcal{T}}(t) - \overline{\mathcal{T}}(t_0) = t - t_0} , \tag{23}$$

d.h. der Erwartungswert des Zeitoperators ist also stets die c-Zahl-Zeit t. Dabei bedeutet $\overline{\mathcal{T}}(t_0) = \mathrm{Spur}\{\mathcal{W}(p, q, t_0)\mathcal{T}(p, q, t_0)\}$.

Schließlich werde noch die Streuung des Zeitoperators

$$\Delta T = \sqrt{\overline{(\mathcal{T} - I\overline{\mathcal{T}})^2}} = \sqrt{\overline{\mathcal{T}^2} - \overline{\mathcal{T}}^2} \tag{24}$$

untersucht. Aus

$$\dot{\overline{\mathcal{T}^2}} = \overline{\overset{0}{\mathcal{T}}\mathcal{T} + \mathcal{T}\overset{0}{\mathcal{T}}} = 2\overline{\mathcal{T}} \tag{25}$$

folgt mit Gl. (23)

$$\overline{\mathcal{T}^2}(t) - \overline{\mathcal{T}^2}(t_0) = \overline{\mathcal{T}}^2(t) - \overline{\mathcal{T}}^2(t_0) ,$$

so daß für die Streuung von $\mathcal{T}$ gilt

$$\Delta T(t) = \Delta T(t_0) . \tag{26}$$

550

Die Streuung des Zeitoperators ist also bei reversiblen Bewegungen stets zeitlich konstant.

Für den zu $\mathcal{H}(p, q, t)$ konjugierten Operator $\mathcal{K}(p, q, t)$ hat man allgemein die Bedingung

$$[\mathcal{H}\mathcal{K}]_- = \frac{\hbar}{i}\mathbf{1}\,. \tag{27}$$

Es gilt also

$$\overset{0}{\mathcal{K}} = \mathbf{1} + \left(\frac{\partial \mathcal{K}}{\partial t}\right)_{p,q} \tag{28}$$

und damit für den Erwartungswert

$$\overline{\mathcal{K}}(t) - \overline{\mathcal{K}}(t_0) = t - t_0 + \int_{t_0}^{t} \left(\frac{\partial \overline{\mathcal{K}}}{\partial t'}\right)_{p,q} \mathrm{d}t'. \tag{29}$$

In den Erwartungswert von $\mathcal{K}$ geht also außer t auch noch seine explizite Zeitabhängigkeit ein. Im zeitunabhängigen Fall fällt diese fort und es ist

$$\mathcal{K}(p,q) = \mathcal{C}(p, q)\,. \tag{30}$$

5. – Diskussion.

Die angestellte Analyse zeigt, daß die Zeit in der Quantenmechanik eine Doppelrolle spielt. Insoweit sie einen reinen Ordnungsparameter darstellt, der a priori zur Beschreibung des *topologischen Nacheinander* vorgegeben ist, kann sie keine (streuende) Observable sein und muß daher als c-Zahl t in die Theorie eingehen. Andererseits ist die Zeit eines quantenmechanischen Systems, die nach (16) durch das *dynamische Verhalten des Systems selbst* gegeben ist, auch eine beobachtbare physikalische Größe, also eine Observable, die den Quantisierungsforderungen gemäß durch einen hermitischen Operator $\mathcal{C}$ beschrieben wird. Während diese zweite Eigenschaft — von der man aber meist nicht spricht — quantentheoretisch unproblematisch ist, scheint die Existenz eines scharfen Zeitparameters zunächst im Rahmen der Quantenmechanik unverständlich zu sein ([7]). Hierauf ist jedoch zu sagen, daß auch makroskopische Systeme, an denen makroskopische Zeitmessungen möglich sind, einer quantenmechanischen Beschreibung zugänglich sind. Die Messung der Observablen « Zeit » an solchen « Uhren » liefert eine physikalische Festlegung des Zeit-

([7]) Vgl. E. SCHRÖDINGER: *Berl. Ber.*, 238 (1931).

parameters t, wodurch seine Existenz auch im Rahmen der Quantenmechanik legitim wird. Die Forderung nach dem Vorhandensein des Begriffs « Nacheinander » in der Quantenmechanik führt also über die Observable « Zeit » zur Möglichkeit einer Bestimmung der Zeitordnung.

Die Tatsache, daß der Erwartungswert des Zeitoperators irgendeines Systems stets dem systemunhabhängigen Zeitordnungsparameter t gleich ist, könnte man dazu ausnützen, die c-Zahl t allgemein durch den Erwartungswert des Zeitoperators sogar eines beliebigen Systems zu ersetzen.

Explizite Zeitabhängigkeiten sind nur durch die Betrachtung von Veränderungen in einer vorgegebenen Zeitordnung definierbar. Demnach kann hier nur der c-Zahl-*Parameter* t eingehen. Hier liegt der Ansatzpunkt für die Ableitung der quantenmechanischen Bewegungsgleichungen, die allein aus der Grundforderung der korrespondenzmäßigen Übersetzung bildunabhängig möglich ist. Nur die expliziten Zeitabhängigkeiten haben hierbei demnach eine tiefergehende Bedeutung. Da die statistische Deutung der Quantenmechanik eine andauernd eindeutige Beschreibbarkeit der physikalischen Zustände erfordert (vgl. (4) und (5)), läßt sich rein deduktiv auch die Bewegungsgleichung (7) des Zustandsvektors bildunabhängig gewinnen. Eine weitere direkte Folgerung ist die ganz allgemeine Gültigkeit des Ehrenfestschen Theorems (9). — Totale Zeitabhängigkeiten von der c-Zahl-Zeit sind nur im Zusammenhang mit einem willkürlich wählbaren Bezugssystem im Hilbert-Raum sinnvoll (z.B. Schrödinger- oder Heisenberg-Bild).

Allein mit Hilfe der *Observablen* « Zeit » läßt sich für den Fall von Systemen, deren Hamilton-Operator und damit auch deren Zeitoperator explizit von der Zeit unabhängig ist, aus der Eigenschaft (17) ([8]) die Unbestimmtheitsrelation zwischen Energie und Zeit begründen. Für Systeme mit explizit zeitabhängigem Hamilton-Operator ist der Zeitoperator $\mathfrak{T}$ im allgemeinen selbst explizit zeitabhängig. Da hier eine direkte Berechung von $\mathfrak{T}$ analog zu (14) und (16) nicht möglich ist und die Eigenschaft (21) allein $\mathfrak{T}$ nicht vollständig bestimmt, ist der Zeitoperator in diesem Fall unendlich vieldeutig. Der Grund dafür liegt in der Tatsache, daß ein explizit zeitabhängiger Hamilton-Operator ein System beschreibt, das entweder in Wechselwirkung mit einem makroskopischen System steht oder auf ein bewegtes Koordinatensystem bezogen ist, so daß in den Zeitoperator ein in weiten Grenzen willkürlicher Anteil der Zeitskala des makroskopischen Systems bzw. der vorgeschriebenen Bewegung des Koordinatensystems eingeht.

([8]) Die von W. Pauli (*Handb. d. Phys.* (S. Flügge), Bd. V/1; S. 60) gegen die Beziehung (17) vorgebrachten Bedenken sind nach unserer Ansicht nicht stichhaltig, weil ein in der Argumentation benützter Satz von Dirac (*Quantenmechanik*, S. 56) über die Eigenwerte konjugierter Operatoren nicht zu gelten braucht, wenn die beim Beweis verwendete unitäre Transformation keinen Automorphismus zwischen den ursprünglichen und den transformierten Eigenfunktionen der Operatoren liefert.

SUPPLEMENTO AL VOLUME XII, SERIE X
DEL NUOVO CIMENTO

N. 1, 1959
2° Trimestre

Note on Symmetry Operations in Quantum Mechanics.

R. HAGEDORN

CERN - Geneva

(ricevuto il 24 Dicembre 1958)

CONTENTS. — **Remark.** – **1. Introduction** 1. Complete set of observables. 2. Superselection rules. – **2. Definition of symmetry.** 1. Active and passive interpretation. 2. Formal definition of symmetry. – **3. Consequences of the existence of a symmetry group.** 1. Existence of a unit-ray transformation. 2. Class of transformations θ leading to the same $\hat{\theta}$. 3. Conservation of the norm. 4. Complete orthonormal systems. 5. The restricted distribution law. Linearly independent elements. 6. The operator function. 7. Superselection subspaces. 8. Continuity. 9. Reduction to unitary and antiunitary transformations θ. 10. Dependence on the group element g. Representation up to a factor.

Remark.

This note does not contain any new result. But the proof that symmetries of quantum mechanical systems give rise to either unitary or antiunitary representations takes a new and simple form. In order to state the problem, it seemed, however, necessary to discuss the implications of symmetries in some detail.

1. – Introduction.

1·1. *Complete set of observables.* – We assume a definite physical system S to be given. The system will be described in a Hilbert space $\mathcal{H}$. Take any observable (self-adjoint operator) A' and construct a set $\mathfrak{A}$ in the following

$$A' \in \mathfrak{A} ,$$

way $A \in \mathfrak{A}$ if and only if A commutes with all other elements of $\mathfrak{A}$. Now if $A \in \mathfrak{A}$; $A' \in \mathfrak{A}$ then

$$AA' \in \mathfrak{A}$$

$$aA + bA' \in \mathfrak{A}\,, \qquad (a,\ b \text{ real}).$$

Hence from the distributive and associative law it follows:

$\mathfrak{A}$ is a maximal abelian ring with real coefficients.

Let $B \notin \mathfrak{A}$. This B defines in the same way an abelian ring $\mathfrak{B}$. In this way we may construct a whole set of such rings $\mathfrak{A}$, $\mathfrak{B}$, $\mathfrak{C}$, $\mathfrak{D}$, ... (*). The set of all observables is not a ring since the product of two self-adjoint operators need not be self-adjoint. All operators of any such ring can be brought simultaneously into diagonal form (spectral representation) and hence can be measured without interference. There is a proof by v. NEUMANN (1) that in such a ring $\mathfrak{A}$ there always exists an element A_0 say, such that all elements of the ring are functions of A_0, namely $A_i = f_i(A_0)$. In principle the measurement of A_0 alone would determine completely the state of the system. This might be technically impossible. In fact, normally one takes one convenient operator A_1, say. If now all other elements of $\mathfrak{A}$ cannot be represented as functions of A_1, one takes another one A_2 and so on, till one has as many operators as are necessary in order to generate the whole ring $\mathfrak{A}$. The set of elements $A_1 \ldots A_n$ is called a complete system of commuting observables. In actual cases, one frequently seems to be able to construct such a system. For every ring $\mathfrak{A}$, $\mathfrak{B}$, $\mathfrak{C}$, $\mathfrak{D}$ ... we assume such a set of generating elements: A_i, B_i ... and call the set of rings $\{\mathfrak{A}, \mathfrak{B}, \mathfrak{C} \ldots\}$ complete, if every observable belongs to at least one of the rings.

1·2. *Superselection rules.* – Obviously the multiples of the unit operator commute with all our observables. If there exists, apart from that, another operator X' commuting with all observables, then the Hilbert space can be decomposed into superselection subspaces. No observable will have matrix elements between different superselection subspaces (that is the reason for the name) (2).

Proof: Together with X', X'^{+} (the hermitian conjugate of X) also commutes, hence $X = \frac{1}{2}(X' + X'^{+})$ is hermitian and commutes also. We assume now that X has a discrete spectral representation:

$$X = \sum_i \lambda_i X_i\ ;$$

(*) These rings overlap of course.

(1) J. v. NEUMANN: *Math. Grundlagen der Quantenmechanik* (Berlin 1932).

(2) C. G. WICK, A. S. WIGHTMAN and E. P. WIGNER: *Phys. Rev.*, **88**, 101 (1952).

with

$$X_i^2 = X_i\,; \qquad X_i X_k = 0 \qquad (i \neq k) \text{ and } X_i^+ = X_i,$$

X_i projects into that subspace $\mathcal{H}_i$ of $\mathcal{H}$, where X has the eigenvalue λ_i. (Always $\lambda_i \neq \lambda_k$, if λ_j occurs n_j times, $\mathcal{H}_j$ has dimension $n_j \leqslant \infty$).

From

$$[XA] = \sum_i \lambda_i [X_i A] = 0$$

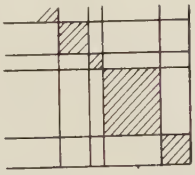
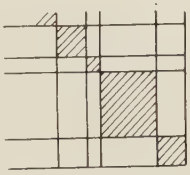
follows by multiplication with X_k to left and right

$$\lambda_k [A X_k] = 0 \qquad \text{or} \qquad [A X_k] = 0 \qquad\qquad \lambda_k \neq 0.$$

Let $\varphi_i \in \mathcal{H}_i$, $\varphi_k \in \mathcal{H}_k$, $i \neq k$. Then at least one of λ_i, λ_k is $\neq 0$, say $\lambda_k \neq 0$. Thus

$$\langle \varphi_i, A\varphi_k \rangle = \langle X_i \varphi_i, A X_k \varphi_k \rangle = \langle X_i \varphi_i, X_k A \varphi_k \rangle = \langle X_k X_i \varphi_i, A\varphi_k \rangle = 0\,.$$

Since A can be any observable, the proof is complete. Any observable can then be brought to the following form:

$A =$  , $B =$  , etc.

and the dimensions (in general infinite) of these boxes on the diagonal will be the same for all observables. Examples of such superselection rules are the charge conservation and baryon conservation.

One sees that in the presence of superselection rules arbitrary unitary transformations cannot be admitted, since they would in general mix up the superselection subspaces. Those unitary transformations, however, which transform from a basis $\{\psi_i\}$ belonging to the ring $\mathfrak{A}$ to a basis $\{\varphi_j\}$ belonging to the ring $\mathfrak{B}$, leave the subspace invariant and are admitted. We may even admit those transformations which furthermore induce a permutation of the superselection subspaces.

2. – Definition of symmetry.

A transformation group G may be defined in many ways, *e.g.*:

— Interchanging of particles,

— Charge conjugation (better perhaps: matter conjugation),

— Lorentz transformations,

— Gauge transformations, etc.

2·1. *Active and passive interpretation.* – There are two ways of interpretation:

i) passive,

ii) active.

Example: The three dimensional rotation may be interpreted passively as a rotation of the co-ordinate system or actively as a bodily rotation of the system S under observation.

In what follows we shall adopt throughout the active interpretation, which is always feasible, whereas the passive one looks very strange, if not impossible in some cases.

2·2. *Formal definition of symmetry.* – Now we define what we mean by saying that our physical system S is invariant under a group G, which we shall then call a symmetry group of S.

A material system S is said to have a symmetry group G, if for every $g \in G$ there exists

i) another material system $S' = gS$ (symbolically),

ii) a definite function F_g for all observables

such that

α) $A' = F_g(A)$ is again an observable, and

β) the possible results of measuring A' in S' are the same as for measuring A in S and both have the same probability distribution.

In order to put this physical definition in a more mathematical form, we introduce a special notation:

A physical state does not determine an element of $\mathcal{H}$, since together with φ also $\omega\varphi$ with $|\omega| = 1$ represents the same physical state.

For reasons of clarity, we shall denote by:

(d1)
f, g, h ... elements of $\mathcal{H}$,

φ, ψ, φ ... state vectors, *i.e.* elements with norm 1,

$\hat{\varphi}, \hat{\psi}, \hat{\varphi}$ ... unit-rays, representing physical states,

$\hat{f}, \hat{g}, \hat{h}$ ... rays.

A ray $\hat{f}$ is the *set* of all elements λf with $0 \leqslant |\lambda| < \infty$ (f fixed). A unit-ray $\hat{\varphi}$ is the *set* of all state vectors $\omega\varphi$ with $|\omega| = 1$ (φ fixed). Rays are not vectors. Addition, etc., are not defined.

558

We now give the formal definition of symmetry:

(D) If we denote the state vectors of S by φ, ψ, ... and those of S' by φ', ψ', ..., then there is a one-to-one correspondence of the physical states (rerpesented by unit-rays)

$$\hat{\varphi} \underset{g}{\longleftrightarrow} \hat{\varphi}'$$

such that the probabilities are conserved:

$$|\langle \varphi', \psi' \rangle| = |\langle \varphi, \psi \rangle|$$

if $\varphi' \in \hat{\varphi}$, $\psi' \in \hat{\psi}'$ and $\varphi \in \hat{\varphi}$, $\psi \in \hat{\psi}$.

We shall quote this definition henceforward as (D).

3. – Consequences of the existence of a symmetry group.

3·1. *Existence of a unit-ray transformation.* – It follows from (D) that the existence of a symmetry group G implies only the existence of a group $\hat{T}$ of unit-ray transformations $\hat{\theta}$, which maps the set of all unit-rays onto itself

$$\hat{\varphi} \longleftrightarrow \hat{\varphi}' = \hat{\theta}\hat{\varphi} .$$

That this transformation group is uniquely defined by G and is isomorphic to G, follows from physical experience. In this form, however, it is useless for quantum mechanics, since the superposition principle is true for state vectors rather than for unit-rays. The problem is therefore to find a group of transformations θ defined for elements of the whole Hilbert space, which is homomorphic, or even possibly isomorphic to the group $\hat{\theta}$, and as simple as possible.

3·2. *Class of transformations θ leading to the same $\hat{\theta}$.* – Let $\hat{T}$ be a group of unit-ray transformations. If $\hat{\theta} \in \hat{T}$, we consider all transformations θ of $\mathcal{H}$ which correspond to $\hat{\theta}$. Correspondence is defined by

(d2)
$$\begin{cases} \theta \sim \hat{\theta} & \text{if} \\ \theta\varphi \in \hat{\theta}\hat{\varphi} & \text{for all} \quad \varphi \in \hat{\varphi} . \end{cases}$$

Consider the set of all non-singular θ, which correspond to all $\hat{\theta} \in \hat{T}$. These θ form a group T, which is homomorphic to the group $\hat{T}$. This is a very large homomorphism, since many θ correspond to the same θ; for instance θ_1 and θ_2 may be defined completely different ly for all elements with norm $\neq 1$ and yet

correspond to the same $\hat{\theta}$. Consider the unit element of $\hat{T}$, $\hat{\theta} = E$. The set of transformations $\theta \sim E$ forms an invariant sub-group T_0 of T (the kernel of the homomorphism). The factor group T/T_0 is then isomorphic to $\hat{T}$. If $\theta \sim \hat{\theta}$, then the coset $\theta T_0 = T_0\theta$ contains *all* transformations which correspond to $\hat{\theta}$; we write $\theta T_0 \sim \hat{\theta}$. The transformations $\theta \in T_0$ have the property of multiplying state vectors φ only by unimodular complex numbers (leaving their « length » and « direction » constant). On other elements with norm $\neq 1$, they may be defined arbitrarily.

Any coset of T_0 may be represented by one of its elements, but if we select one element of each coset, then in general these elements will not form a group.

It will be shown that elements θ can be selected from the cosets such that they are either unitary or antiunitary and form a representation up to a factor.

3·3. *Conservation of the norm.* – First we postulate

$$\theta f = \|f\| \cdot \theta \frac{f}{\|f\|} . \tag{3.1}$$

By this θ is defined in the whole of $\mathcal{H}$ once it is defined for state vectors. These θ still form a group. The invariant subgroup T_0 now contains all those θ_0 which merely multiply each element of $\mathcal{H}$ with a unimodular complex factor such that $\theta_0 f = \omega_{f/\|f\|} \cdot f$. The factor depends only on $f/\|f\|$. We shall henceforward call this invariant sub-group Ω. This is in fact a very strong restriction. Its effect is that two transformations θ and θ', which both correspond to one single $\hat{\theta}$, differ at most by a $\theta_0 \in \Omega$.

Obviously the restriction is not too strong, since it does not affect the transformation of state vectors.

3·4. *Complete orthonormal systems.*

Theorem 1: Complete orthonormal systems are transformed into complete orthonormal systems.

Proof: Let $\{\psi_i\}$ be a complete orthonormal system, and

$$\theta \sim \hat{\theta}$$

$$\theta' \sim \hat{\theta}^{-1}$$

then $\theta\theta' \sim E$. $\psi_i \to \psi_i' = \theta\psi_i$ and from (D)

$$|\langle \theta\psi_i, \theta\psi_k \rangle| = |\langle \psi_i, \psi_k \rangle| = \delta_{ik} . \tag{3.2}$$

It remains to show that $\{\psi_i'\}$ is complete. Assume it is not; then there exists a ψ_0' orthogonal to all ψ_i'. Then $\theta'\psi_0' = \psi_0$ would be orthogonal to all ψ_i against the assumption.

3·5. *The restricted distribution law. Linearly independent elements.*

Theorem 2:

i) If f, g, h ... are linearly independent then θf, θg, θh ... are also.

ii) $\theta(f+g) = \omega(f, g)\theta f + \omega(g, f)\theta g$ with $|\omega(f, g)| = 1$ for any pair of elements f, g.

Proof:

i) By induction. Let it be true for n vectors $f_1 \dots f_n$.

$$f_1 \dots f_n \quad \text{span a subspace} \quad M_n ,$$
$$\theta f_1 \dots \theta f_n \quad \text{span a subspace} \quad M_n^\theta .$$

Let $f_{n+1} = f$ be linearly independent of $f_1 \dots f_n$. Then there must exist a h such that

$$\langle h, f_i \rangle = 0 ; \qquad i = 1 \dots n,$$
$$\langle h, f \rangle \neq 0 .$$

It follows that

$$\langle \theta h, \theta f_i \rangle = 0 ; \qquad i = 1 \dots n,$$
$$\langle \theta f, \theta f \rangle \neq 0 .$$

Therefore θf is linearly independent of $\theta f_1 \dots \theta f_n$. Since the Theorem is true for $n = 1$, it holds for all n.

ii) For $f = \lambda g$ it is trivial. Let f, g be linearly independent. They span a sub-space M. For any $h \in M^\perp$

$$|\langle \theta h, \theta(f+g) \rangle| = |\langle h, f+g \rangle| = 0 .$$

Hence $\theta(f+g)$ belongs to the sub-space spanned by θf and θg:

$$\theta(f+g) = \lambda(f, g)\theta f + \mu(f, g)\theta g .$$

Define

$$k(f, g) = f - \frac{\langle g, f \rangle}{\| g \|^2} g ; \qquad \langle k, g \rangle = 0 .$$

Then

$$|\langle \theta k, \theta(f+g)\rangle| = |\langle k, f+g\rangle| = |\langle k, f\rangle|$$

but also

$$|\langle \theta k, \theta(f+g)\rangle| = |\langle \theta k, \lambda(f,g)\theta f\rangle| = |\lambda(f,g)|\cdot|\langle k, f\rangle|.$$

Hence $|\lambda(f,g)| = 1$. From interchanging f and g it follows that $\lambda(f,g) = \mu(g,f) = \omega(f,g)$, *q.e.d.*

In particular:

$$\theta \sum c_i \psi_i = \sum c_i' \theta \psi_i \qquad \text{with} \qquad |c_i'| = |c_i|. \tag{3.3}$$

3·6. *The operator function.* – We can now write down explicitly the operator function

$$A' = F_g'(A).$$

Let $\theta \sim \hat{\theta}(g)$. If

$$A = \sum a_i |\psi_i\rangle\langle\psi_i|,$$

then

$$F_g'(A) = A' = \sum a_i |\theta\psi_i\rangle\langle\theta\psi_i| = \sum a_i |\psi_i'\rangle\langle\psi_i'| \,. \tag{3.4}$$

We have to show that this is unique. Thus any $\theta \sim \hat{\theta}(g)$ must give the same result. Take $\theta' = \theta_0\theta$; $\theta_0 \in \Omega$:

$$A'' = \sum a_i |\theta'\psi_i\rangle\langle\theta'\psi_i| = \sum c_i |\omega_i\psi_i'\rangle\langle\omega_i\psi_i'| = \sum a_i\omega_i\omega_i^* |\psi_i'\rangle\langle\psi_i'| = A' \,.$$

It follows that A' has the same eigenvalues as A and the transformed state vectors ψ_i' are the new eigenvectors.

Theorem 3: The expectation values are conserved.

Proof:

$$\varphi = \sum c_i \psi_i \,; \qquad \theta\varphi = \sum c_i' \theta\psi_i \,,$$

then

$$\langle \theta\varphi, A'\theta\varphi\rangle = \sum c_j'^* a_i c_i' \langle\theta\psi_j|\theta\psi_i\rangle\langle\theta\psi_i|\theta\psi_e\rangle = \sum a_i |c_i'|^2 \,,$$

$$\langle \varphi, A\varphi\rangle \quad = \sum a_i |c_i|^2 \,.$$

These are equal on account of (3.3). This means: Corresponding measurements in corresponding states will always yield the same results. Hence the whole physics is unchanged. That is the meaning of our definition of symmetry.

3·7. *Superselection subspaces.*

Theorem 4: All $\theta \in T$ will, apart from a mapping of superselection subspaces onto themselves, at most induce a permutation of these spaces.

Proof:

i) From Theorem 1 it follows that orthogonal subspaces are transformed into orthogonal subspaces.

ii) From Theorem 3 it follows that the transformed observables have no matrix elements between the transformed superselection subspaces.

iii) The transformation of the observables is a mapping of the set of of all observables onto itself. This set defines the decomposition $\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_2 \oplus \mathcal{H}_3 \oplus \ldots$ up to the order in which the $\mathcal{H}_i$ are counted, thus a a certain $\theta \in T$ induces a well-defined permutation of the $\mathcal{H}_i$ (which might be the unit-permutation).

3·8. *Continuity.* – There are two types of continuity:

i) By symmetry operations on the physical system neighbouring physical states are transformed into neighbouring physical states. This is a physical experience. It is contained in our definition of symmetry (D):

$$|\langle \psi', \varphi' \rangle| = |\langle \psi, \varphi \rangle| .$$

Let $\varphi = \varphi(t)$ follow a continuous path, then both sides have to be continuous functions of t. This does not imply that $\varphi'(t)$ is also continuous, since it may contain any discontinuous phase factor $\omega(t)$.

ii) The symmetry group G may be continuous or partly continuous: There may exist continuous paths $g(t)$.

We shall presently discuss the first type of continuity. For this we define distance and neighbourhood of unit-rays:

(d3) The distance of two unit-rays $\hat{\varphi}$ and $\hat{\psi}$ is $\hat{d} = \min_\omega \|\varphi - \omega\psi\|$ where $\varphi \in \hat{\varphi}$ and $\psi \in \hat{\psi}$ is an otherwise arbitrary pair of state vectors.

(d4) *A* ε-neighbourhood of $\hat{\varphi}$ is the set of all unit-rays $\hat{\psi}$ whose distance from $\hat{\varphi}$ is $< \varepsilon$.

With these definitions the first type of continuity can be stated as.

Theorem 5: The transformations $\hat{\theta}$, which correspond to elements $g \in G$ are continuous.

This is obvious since neighbouring rays are transformed in neighbouring rays.

Theorem 6: In the set of all $\theta \sim \hat{\theta}$ there is at least one which is continuous in the whole considered superselection subspace.

Sketch of a proof: By construction. Take an arbitrary state vector φ_0. It belongs to $\hat{\varphi}$. Take now an arbitrary state vector $\varphi_0' \in \hat{\varphi}'$, where $\hat{\theta}\hat{\varphi} = \hat{\varphi}'$. Define $\theta\lambda\varphi_0 = \lambda\varphi_0'$ for all λ $(0 \leqslant |\lambda| < \infty)$. Hereby θ is defined for multiples of φ_0. We now define it for any element g:

Let ψ be any state vector and $\hat{\psi}$ its unit-ray. Find ω such that:

$$\|\varphi_0 - \omega\psi\| = \min\,; \qquad (\psi_0 = \omega\psi \text{ is nearest to } \varphi_0).$$

Find further ω' such that, with $\psi' \in \hat{\theta}\hat{\psi}$

$$\|\varphi_0' - \omega'\psi'\| = \min\,; \qquad (\psi_0' = \omega'\psi' \text{ is nearest to } \varphi_0').$$

Define $\theta\mu\psi_0 = \mu\psi_0'$ for any μ $(0 \leqslant |\mu| < \infty)$. We have to show that this θ is continuous. Given f, we must find a neighbourhood N_f of f such that for all $g \in N_f$ $\|\theta f - \theta g\| < \varepsilon$.

Put $f = \lambda\psi_0$, $g = \mu\Phi_0$, where ψ_0 and Φ_0 are nearest to φ_0 in the above sense. Then

$$\|\theta f - \theta g\| = \|\lambda\psi_0' - \mu\Phi_0'\| \leqslant |\lambda| \cdot \|\psi_0' - \Phi_0'\| + |\lambda - \mu|\,.$$

To make this smaller than ε, choose

$$|\lambda - \mu| < \varepsilon/2\,,$$

$$\|\Phi_0' - \psi_0'\| < \frac{\varepsilon}{2|\lambda|}\,.$$

To achieve the last inequality, take the set of all Φ_0' fulfilling it. To each of them belongs a unit-ray Φ'. Call the set of these unit-rays $\{\Phi'\}_\delta$. Transform each of these unit-rays with $\hat{\theta}^{-1}$. This gives a set $\{\Phi\}_\delta$. Find the nearest element Φ_0 to φ_0 of each $\Phi \in \{\hat{\Phi}\}_\delta$. These Φ_0 form an open set containing ψ_0 and defining a neighbourhood of ψ_0. Any $g = \mu\Phi_0$ for which Φ_0 lies in this neighbourhood and for which $|\lambda - \mu| < \varepsilon/2$, fulfils the condition.

3·9. *Reduction to unitary and antiunitary transformations θ.*

Theorem 7: Given any $\hat{\theta}$, we can choose such a $\theta \sim \hat{\theta}$ that θ is either unitary or antiunitary in the whole considered superselection subspace.

This theorem was proved by E. P. Wigner [3] without taking into account the antiunitary operators. This has been done explicitly also by E. P. Wigner

[3] E. P. Wigner: *Gruppentheorie und ihre Anwendungen auf die Quantenmechanik* (Brunswick, 1931), p. 251.

562

in his Lorentz-chair lectures, University of Leyden 1957 (not published). His proof is rather complicated and the following simple proof is the only essentially new thing in this paper.

Proof: The theorem means that given a $\hat{\theta}$ and a $\theta' \sim \hat{\theta}$, we can find such a $\theta_0 \in \Omega$, that $\theta_0\theta' = \theta$ is either unitary or antiunitary. We construct θ_0.

α) Construct a continuous $\theta' \sim \hat{\theta}$. following the lines of the proof of Theorem 6. From Theorem 2 it follows that for linearly independent elements f, g, h

$$\theta'(f+h+g) = \omega(f+h, g)[\omega(f, h)\theta' f + \omega(h, f)\theta' h] + \omega(g, f+h)\theta' g \,.$$

Interchanging g and h and comparing the coefficients leads to

$$\begin{aligned} &\omega(f+g, h)\omega(f, g) = \omega(f+h, g)\omega(f, h) && \text{coefficient of } \theta' f, \\ &\omega(f+g, h)\omega(g, f) = \omega(g, f+h) && \text{coefficient of } \theta' g, \\ &\omega(f+h, g)\omega(h, f) = \omega(h, f+g) && \text{coefficient of } \theta' h. \end{aligned}$$

This gives, by dividing the first two and eliminating $\omega(f+h, g)$ with the help of the third equation

$$\omega(f, g) = \frac{\omega(h, f+g)}{\omega(h, f)} \cdot \omega(f, h) \cdot \frac{\omega(g, f)}{\omega(g, f+h)} \,. \tag{3.5}$$

From $\theta'(f+0) = \omega(f, 0)\theta' f = \theta' f$ follows $\omega(f, 0) = 1$. We put $\omega(0, f) = 1/u(f)$ (*). Now letting $h \to 0$ one obtains

$$\omega(f, g) = \frac{u(f)}{u(f+g)} \,.$$

$u(f)$ is unimodular by definition. That it is continuous and independent of the norm $\|f\|$, follows from the corresponding properties of θ'.

We have therefore a unimodular function $u(f)$ such that

$$u(f+g)\theta'(f+g) = u(f)\theta' f + u(g)\theta' g \,.$$

$u(f)$ is an element $\theta_0 \in \Omega$. Now $\theta_0\theta' = \theta$ is distributive: $\theta(f+g) = \theta f + \theta g$, and θ is again continuous.

β) (4) $\quad \langle \theta(f+g), \theta(f+g) \rangle = \langle f+g, f+g \rangle = \|f\| + \|g\| + 2 \operatorname{Re} \langle f, g \rangle \,,$

$$\langle \theta(f+g), \theta(f+g) \rangle = \|f\| + \|g\| + 2 \operatorname{Re} \langle \theta f, \theta g \rangle \,.$$

(4) The following argument is due to Dr. J. M. Jauch (private communication).

(*) This is not as obvious as it first seems. See the discussion in the appendix, p. 85.

Hence

$$\mathrm{Re}\,\langle \theta f, \theta g\rangle = \mathrm{Re}\,\langle f, g$$

and from (D)

$$|\langle \theta f, \theta g\rangle| = |\langle f, g\rangle|.$$

Only two solutions exist: $\mathrm{Im}\,\langle \theta f, \theta g\rangle = \pm\,\mathrm{Im}\,\langle fg\rangle$ thus:

Either	$\langle \theta f, \theta g\rangle = \langle f, g\rangle$	θ is unitary,
or	$\langle \theta f, \theta g\rangle = \langle g, f\rangle$	θ is antiunitary

Since θ is continuous, it is either unitary in the whole subspace or antiunitary in the whole subspace. Note that the same cannot be concluded for the whole of $\mathcal{H}$.

3·10. *Dependence on the group element g. Representation up to a factor.* – We now have unitary or antiunitary θ's. For this we have paid heavily. Beginning with a given θ', we found the unimodular $u(f)$ uniquely defined up to a unimodular factor $\omega(\theta)$, which no longer depends on f.

Assume we had started with a θ'', we would have found a $u''(f)$ represented by the group element $\theta_0'' \in \Omega$. Now $\theta_0\theta'$ and $\theta_0''\theta''$ can differ only by another $\theta_0''' \in \Omega$. But this θ_0''' can only be a constant unimodular ω, since otherwise θ_0 and θ_0'' were *not* uniquely defined (apart from a constant factor).

Therefore we can postulate that θ shall be unitary or antiunitary (we cannot choose between these two) but we pay for this by loosing all freedom but for the freedom of multiplying θ by a unimodular $\omega(\theta)$. Apart from that, θ is then a uniquely defined function of $\hat{\theta}$ and therefore of the group element g.

With a choice of the remaining free function $\omega(\theta)$ everything is fixed. We have then chosen exactly one transformation $\theta(g)$ from every coset. We cannot expect these transformations to form a group, but we have always

$$\theta(g_1)\theta(g_2) = \omega(g_1 g_2)\theta(g_1 g_2)\,.$$

This is called a « representation up to a factor ».

It should be clear that if we admit arbitrary phase factors, then the θ form a group. Its invariant sub-group corresponding to $\hat{\theta} = E$ is now the group ω of complex unimodular numbers, and the factor group T/ω is still isomorphic to the group G. It is by fixing these phase factors that the group property is lost. It may happen, however, that these phase factors can be chosen such that the θ's still form a group. Under which circumstances this is possible, is discussed in full generality by V. BARGMANN (*Ann. o. Math.*, **59**, 1 (1954)).

561

* * *

The author wishes to thank Professor J. M. JAUCH, who is visiting CERN, for many stimulating and clarifying discussions. He is indebted to Dr. H. A. TOLHOEK for discussions and for giving him a copy of notes on the Lorentz-chair lectures of Professor E. P. WIGNER. The notes were written, in great detail, by Dr. R. M. F. HOUTAPPEL, and include additional calculations by him. Finally, thanks are due to many members of the Theoretical Group of CERN.

APPENDIX

In the proof of Theorem 7 we put $\omega(0, f) = 1/u(f)$.

One may wonder whether $\omega(0, f)$ makes sense, since from

$$\theta'(f+h) = \omega(f, h)\theta' f + \omega(h, f)\theta' h$$

no information is obtained regarding $\omega(0, f)$ and even the $\lim_{h\to 0} \omega(h, f)$ cannot be found from it.

That such a limit exists if $h \to 0$ in a well defined way, is shown as follows:

From the proof of Theorem 6 it follows that $\theta' \lambda f = \lambda \theta' f$ for any λ. From Theorem 2 ii) it follows that

$$\theta'(f+\lambda f) = (1+\lambda)\theta' f = [\omega(f, \lambda f) + \lambda\omega(\lambda f, f)]\theta' f\,,$$

hence

$$\lambda = -\frac{1-\omega(f, \lambda f)}{1-\omega(\lambda f, f)}\,.$$

Since $|\omega| = 1$, the numerator and denominator of the ratio lie both on a unit circle with centre at $+1$ in the complex plane. In general the ratio and with it λ will be complex. Real λ are possible only for $\omega(f, \lambda f) = \omega(\lambda f, f)$ for $\neq 1$ this gives $\lambda = -1$. Other real λ are possible only for $\omega(f, \lambda f) \equiv -\omega(\lambda f, f) = 1$. This is also a trivial solution for complex λ. Any solution for complex λ must coincide with this solution when λ becomes real.

From (3.5) follows that the limit

$$\lim_{0\to h} \frac{\omega(h, f+g)}{\omega(h, f)} = \omega(f, g)\,,$$

exists independently of how $h \to 0$. Since g is arbitrary, $f+g$ is arbitrary.

Therefore with $f+g=k$

$$\lim_{h\to 0}\frac{\omega(h,k)}{\omega(h,f)} \text{ exists independentlu of how } h\to 0\,.$$

Choose now $h=\lambda\cdot k_0$; where $k_0=k/\|k\|$ and take λ real. Then

$$\omega(\lambda k_0, k)=1$$

for all $\lambda>0$ and it follows that

$$\lim_{h\to 0}\frac{\omega(h,k)}{\omega(h,f)}=\lim_{\lambda\to 0}\frac{1}{\omega(\lambda k_0,f)}=u_{k_0}(f)$$

exists. Since k_0 is arbitrary, we conclude that

$$\lim_{\|h\|\to 0}\omega(h,f)=\frac{1}{u_{h_0}(f)}\,.$$

We have then

$$\omega(f\cdot g)=\lim_{h\to 0}\frac{\omega(h\,f+g)}{\omega(h,f)}=\frac{\lim_{\|h\|\to 0}\omega(h,f+g)}{\lim_{\|h\|\to 0}\omega(h,f)}=\frac{u_{h_0}(f)}{u_{h_0}(f+g)}\,.$$

Since this is independent of h_0, it follows by comparing it with another limit using k_0', that

$$\frac{u_{h_0}(f)}{u_{h_0'}(f)}=\frac{u_{h_0}(k)}{u_{h_0'}(k)}=\ldots=F(h_0,h_0')\,.$$

Therefore $u_{h_0}(f)$ and $u_{h_0'}(f)$ differ only by a unimodular factor which does not depend on f. Since this is irrelevant for the further arguments, we may simply write $\omega(0,f)=1/u(f)$ as was done in the proof of Theorem 7.

One may wonder how $\theta'\lambda f=\lambda\theta' f$ for arbitrary complex λ could be true, since we finally arrive at a θ which may be anti-linear. But obviously this property has gone then:

$$\theta\lambda f=u(\lambda f)\theta'\lambda f=\lambda\cdot u(\lambda f)\theta' f \quad \text{and it may be that} \quad u(\lambda f)\neq u(f)\,.$$

In a way the property of θ':

$$\theta'\lambda f=\lambda\theta' f$$

and the property of θ:

$$\theta(f+g)=\theta f+\theta g$$

are opposite extremes, which coincide only for unitary θ.

566

SUPPLEMENTO AL VOLUME XII, SERIE X
DEL NUOVO CIMENTO

N. 1, 1959
2° Trimestre

Two Theorems on Scattering (*).

I. I. ZINNES

University of Oklahoma - Norman, Okla.

(ricevuto il 19 Gennaio 1959)

CONTENTS. — Introduction. – **1.** The evanescence of wave packets. **2.** Uniqueness of the free hamiltonians of a scattering system.

Introduction.

In this paper we shall prove two theorems which belong to the still growing body of literature on the formal theory of scattering (1). Any wave packet representing a free particle develops in time in such a manner that the probability of finding the particle in any finite region of space vanishes in the limit as $t \to \pm \infty$ (Theorem 1). 2) In a multi-channel scattering process the set of (free) Hamiltonians defining the asymptotic motion of the fragments is unique (Theorem 2). Although Theorem 1, called for obvious reasons the theorem on the evanescence of wave packets, is well known, a rigorous proof (1) does not appear to be in the literature. In Theorem 2 we assume the appropriateness of Jauch's definition (2) of a multichannel scattering process, a definition which is an extension of his previous definition (3) of a single scat-

(*) Supported by the National Science Foundation.

(1) L. I. SCHIFF: *Quantum Mechanics* (New York, 1949). A particular case of the Theorem I referring to a one-dimensional minimum wave packet is considered on pp. 54-59.

(2) J. M. JAUCH: *Helv. Phys. Acta* (in print).

(3) J. M. JAUCH: *Helv. Phys. Acta*, **31**, 127 (1958).

tering process. Although these two theorems appear to be unrelated, it is convenient to consider them together, since the proofs are based on very similar mathematical results. This similarity is due to the fact that both theorems are formulated in terms of essentially the same kinetic energy operator, $c\boldsymbol{P}^2$, where $c^2 > 0$.

The proofs of the theorems are based on the theory of linear operators defined on a Hilbert space $\mathcal{H}$ of (square integrable or L^2) state functions. The delta function and non-normalizable functions are thereby completely avoided (1). The principal symbols used are: $\| \ \|$, norm of an element in $\mathcal{H}$ and of a bounded operator (defined everywhere in $\mathcal{H}$); $\rightarrow$, strong convergence; $\overset{w}{\rightarrow}$ weak convergence; and (5) $\overset{u}{\rightarrow}$, convergence in the usual sense of point-wise convergence.

1. – The evanescence of wave packets.

It is perhaps physically reasonable to expect that a wave packet which represents a free particle will develop in such a way that the part of the packet which differs significantly from zero at $t = 0$ will in the course of time fill all space. Because the normalization of the packet is conserved, it is clear that the probability of finding the associated particle in any finite portion of space should converge to zero as $t \rightarrow \infty$. Before stating Theorem 1 which justifies this expectation, we introduce the following notation.

Let $\mathcal{H}$ be an abstract space; $\mathcal{H}(\boldsymbol{X})$, a realization of $\mathcal{H}$ formed by the set of L^2-functions, $f(\boldsymbol{X})$, defined over the infinite three-dimensional Euclidean space E_3 of vectors $\boldsymbol{X} = (x, y, z)$; and $\mathcal{H}(\boldsymbol{K})$, a realization of $\mathcal{H}$ formed by the Fourier transforms (also a set of L^2-functions) of the elements of $\mathcal{H}(\boldsymbol{X})$. If the abstract element $f \in \mathcal{H}$ is represented by $f(\boldsymbol{X}) \in \mathcal{H}(\boldsymbol{X})$ and by $f(\boldsymbol{X}) \in \mathcal{H}(\boldsymbol{K})$, then the momentum operator $\boldsymbol{P}$ (with $\hbar = 1$) has a simple representation (6) in $\mathcal{H}(\boldsymbol{K})$:

$$(\boldsymbol{P}^2 f)(\boldsymbol{K}) = \boldsymbol{K}^2 f(\boldsymbol{K}) \tag{1}$$

(4) An outline of the mathematical theory will be found in reference (3). More detailed expositions will be found in B. VON SZ.-NAGY: *Ergebnisse der Mathematik und Ihrer Grenzgebiete*, **5**, 5 (1942); F. RIESZ and B. VON SZ.-NAGY: *Functional Analysis* (New York, 1955).

(5) This symbol is not standard and is used in this paper for the sake of clarity. Normally the symbol for strong convergence is also used for point-wise convergence because its meaning is rarely ambiguous in context.

(6) J. VON NEUMANN: *Math. Ann.*, **104**, 570 (1931). It may not be superfluous to point out here that the terms wave packet and element of a Hilbert space are synonymous.

provided $[\boldsymbol{K}^2 f(\boldsymbol{K})]^2$ is integrable, *i.e.*, provided $\boldsymbol{K}^2 f(\boldsymbol{K}) \in \mathcal{H}(\boldsymbol{K})$. Also, if t is a real variable and $\omega(t)$ is the limit of a sequence of polynomials, then

$$(\omega(\boldsymbol{P}^2)f)(\boldsymbol{K}) = \omega(\boldsymbol{K}^2)f(\boldsymbol{K})\,, \tag{2}$$

provided again $\omega(\boldsymbol{K}^2)f(\boldsymbol{K}) \in \mathcal{H}(\boldsymbol{K})$.

Let Δ be an arbitrary Borel set of points of E_3 with finite Lebesgue measure $\mu(\Delta)$. Also, let E_Δ be a projection operator, which when applied to $f(\boldsymbol{X}) \in \mathcal{H}(\boldsymbol{X})$, has the following property:

$$(E_\Delta f)(\boldsymbol{X}) = \begin{cases} f(\boldsymbol{X}), & \text{if} \quad \boldsymbol{X} \in \Delta\,, \\ 0\,, & \text{otherwise}\,. \end{cases} \tag{3}$$

Therefore, if a particle is represented by the normalized wave packet $f_t = \psi_t/\|\psi_t\|$ at time t, the probability of finding it in Δ at time t is $\|E_\Delta \psi_t\|^2 / \|\psi_t\|^2$, or, if U_t is a unitary group and $\psi_t = U_t\psi$, then this probability is given by $\|E_\Delta U_t\psi\|^2/\|\psi\|^2$. The theorem to be proved may now be stated in the following way.

THEOREM I. Let ψ be an arbitrary element of $\mathcal{H}$; $U_t = \exp[-iH_0 t]$; $H_0 = c\boldsymbol{P}^2$, where $c^2 > 0$. Then $\|E_\Delta U_t\psi\| \to 0$ as $t \to \pm\infty$. Although it is assumed that the non-relativistic Hamiltonian describes the free motion of the particle whose wave packet at $t = 0$ is $\psi(\boldsymbol{X})$, it will be clear that the proof can readily be extended so that the theorem refers to a wider class of Hamiltonians, one which includes the relativistic Hamiltonian $\sqrt{m^2 + \boldsymbol{P}^2}$, for example.

The theorem will be proved with the aid of six lemmas. The first four establish the validity of the theorem for a set of elements dense in $\mathcal{H}$. With the aid of the last two it is shown that the theorem holds for all elements of $\mathcal{H}$, *i.e.*, for every wave packet. Since the proof of the first lemma is readily available, it is stated without proof.

Lemma 1. (Wiener's Displacement Theorem) [7]. If $\varphi(\boldsymbol{X})$ belongs to L^2 and if almost everywhere its Fourier transform differs from zero, then if $\psi(\boldsymbol{X})$ belongs to L^2 and $\varepsilon > 0$, there exists an integer N together with a set of real vectors $\boldsymbol{\Lambda}_n$ and complex numbers A_n $(n = 1, 2, \ldots, N)$ such that

$$\int\int\int_{-\infty}^{\infty} \left| \psi(\boldsymbol{X}) - \sum_{n=1}^{N} A_n \varphi(\boldsymbol{X} - \boldsymbol{\Lambda}_n) \right|^2 \mathrm{d}\boldsymbol{X} < \varepsilon\,,$$

where $\mathrm{d}\boldsymbol{X} = \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$. In this form, Lemma 1 is a slight generalization of the displacement theorem as stated in reference [7], where the theorem refers to

[7] N. WIENER: *The Fourier Integral* (Cambridge, 1933), p. 100.

functions defined on the line. However, it is clear that by considering the closed linear manifold formed by products of three (or more) functions, each defined on the line, Wiener's theorem must hold also for L^2-functions defined on a space of three (or more) dimensions.

For the generator φ of the displacement theorem we may use $\varphi(\boldsymbol{K}) = \exp[-a^2\boldsymbol{K}^2]$ with $a^2 > 0$ and fixed, since (1) its Fourier transform vanishes nowhere in E_3 and (2)

$$\|\varphi\|^2 = \iiint \exp[-2a^2\boldsymbol{K}^2]\,\mathrm{d}\boldsymbol{K} = \left(\frac{\pi}{2a^2}\right)^{\frac{3}{2}},$$

i.e., $\varphi \in \mathcal{H}$. If we define

$$\varphi_{\boldsymbol{K}_0}(\boldsymbol{K}) = \varphi(\boldsymbol{K} - \boldsymbol{K}_0) = \exp[-a^2(\boldsymbol{K} - \boldsymbol{K}_0)^2], \tag{4}$$

we readily deduce that $\varphi_{\boldsymbol{K}_0} \in \mathcal{H}$ and that the Fourier transform of $U_t\varphi_{\boldsymbol{K}_0})(\boldsymbol{K})$ is

$$(U_t\varphi_{\boldsymbol{K}_0})(\boldsymbol{X}) = (\tfrac{1}{2}\pi)^{\frac{3}{2}} \int \exp[-a^2(\boldsymbol{K} - \boldsymbol{K}_0)^2 - ictK^2 + i\boldsymbol{K}\cdot\boldsymbol{X}]\,\mathrm{d}\boldsymbol{K} = \tag{5}$$

$$= [2(a^2 + ict)]^{-\frac{3}{2}} \exp[-i\boldsymbol{K}_0]\cdot(ct\boldsymbol{K}_0 - \boldsymbol{X})] \exp -\left[\frac{(ct\boldsymbol{K}_0 - \frac{1}{2}\boldsymbol{X})^2}{a^2 + ict}\right],$$

and that

$$|(U_t\varphi_{\boldsymbol{K}_0})(\boldsymbol{X})|^2 = [4(a^4 + c^2t^2)]^{-\frac{3}{2}} \exp\left[-\frac{2a^2(\frac{1}{2}\boldsymbol{X} - ct\boldsymbol{K}_0)^2}{a^4 + c^2t^2}\right]. \tag{6}$$

This result leads to

Lemma 2 (8). $|(U_t\varphi_{\boldsymbol{K}_0})(\boldsymbol{X})|^2 \leqslant [4(a^4 + c^2t^2]^{-\frac{3}{2}}$.

This lemma together with (3) implies that

$$\|E_\Delta U_t\varphi_{\boldsymbol{K}_0}\|^2 = \int_\Delta |(U_t\varphi_{\boldsymbol{K}_0})(\boldsymbol{X})|^2\,\mathrm{d}\boldsymbol{X} \leqslant [4(a^4 + c^2t^2)]^{-\frac{3}{2}} \int_\Delta \mathrm{d}\boldsymbol{X} = \mu(\Delta)[4(a^4 + c^2t^2)]^{-\frac{3}{2}}. \tag{7}$$

Therefore we obtain

Lemma 3. $\|E_\Delta U_t\varphi_{\boldsymbol{K}_0}\| \overset{u}{\to} 0$ as $t \to \pm\infty$ (even uniformly in $\boldsymbol{K}_0$).

The result is exactly Theorem 1 when the latter is applied to the set of

(8) This result was obtained in another context by J. M. Cook: *Journ. Math. Phys.*, **36**, 82 (1957).

elements $\{\varphi_{K_0}\}$ in which all vectors K_0 are represented. In addition, since the operators $\{E_\Delta U_t\}$ are linear and the manifold $\mathcal{D}$ formed from the set $\{\varphi_{K_0}\}$ is dense in $\mathcal{H}$ according to Lemma 1, we have verified

Lemma 4. A linear manifold $\mathcal{D}$ dense in $\mathcal{H}$ exists such that for all $f \in \mathcal{D}$, $\|E_\Delta U_t f\| \xrightarrow{u} 0$ as $t \to \pm\infty$.

If an arbitrary sequence $\{t_n\}$ is selected from the interval $0 \leqslant t < \infty$ so that $t_n \to \infty$ as $n \to \infty$, or if a similar sequence is selected from the interval $-\infty < t \leqslant 0$ so that $t_n \to -\infty$ as $n \to \infty$, then Lemma 4 implies the

Corollary. For elements in the domain of Lemma 4, $\|E_\Delta U_{t_n} f\| \xrightarrow{u} 0$ if $t_n \to \pm\infty$ as $n \to \infty$.

To complete the proof of the main theorem we must show that Lemma 4 is valid for *all* elements in $\mathcal{H}$, *i.e.*, for each element in $\overline{\mathcal{D}}$, the closure of $\mathcal{D}$. This extension of Lemma 4 can be carried out with the aid of the following two lemmas. The first of these is the Banach-Steinhaus Theorem (9). Since this theorem is one of the basic theorems in the theory of linear operators, it will not only be stated here in a form useful for the present demonstration, but a relatively transparent proof will also be given.

Lemma 5. Let $\{A_n\}$ be a sequence of linear bounded operators such that

1) $\|A_n\| \leqslant M$ for all n

2) $A_n f$ converges strongly as $n \to \infty$ for each f in a domain $\mathcal{D}$, which is dense in $\mathcal{H}$.

Then $A_n \to A$ on all $\mathcal{H}$, where A is a bounded operator and $\|A\| \leqslant M$.

Proof. Since the Lemma is trivial if $M = 0$, we shall assume that $M > 0$.

1) To prove that $\{A_n\}$ converges strongly in $\mathcal{H}$, it must be shown that given any $\varepsilon > 0$ there exists an N such that for any $g \in \mathcal{H}$

$$\|A_m g - A_{m'} g\| < \varepsilon ; \qquad m,\ m' > N.$$

Since $\mathcal{D}$ is dense in $\mathcal{H}$, corresponding to any $g \in \mathcal{H}$, there exists a sequence $\{g_n\}$ in $\mathcal{D}$ such that $g_n \to g$ as $n \to \infty$. Therefore

$$\begin{aligned}\|A_m g - A_{m'} g\| = \|(A_m - A_{m'})(g_n - g + g_n)\| \leqslant \\ \leqslant \|(A_m - A_{m'})(g_n - g)\| + \|(A_m - A_{m'})g_n\| \leqslant \\ \leqslant 2M\|g_n - g\| + \|(A_m - A_{m'})g_n\| .\end{aligned}$$

(9) E. Hille: *Functional Analysis and Semi-Groups*, vol. **31**, in *Amer. Math. Soc. Colloq. Pub., Amer. Math. Soc.* (New York, 1948).

Since $g_n \to g$, N_1 exists such that $\|g_n - g\| < \varepsilon/4M$ for $n > N_1$. Also, since $g_n \in \mathcal{D}$ and $\{A_m g_n\}$ converges strongly as $m \to \infty$ by hypothesis, N_2 exists such that $\|(A_m - A_{m'})g_n\| \leqslant \varepsilon/2$ for $m, m' > N_2$. Therefore

$$\|A_m g - A_{m'} g\| \leqslant 2M \cdot \frac{\varepsilon}{4M} + \frac{\varepsilon}{2} = \varepsilon ,$$

for $m, m' > \mathrm{Max}(N_1, N_2)$, and $\{A_n g\}$ is a Cauchy sequence for arbitrary $g \in \mathcal{H}$. Since $\mathcal{H}$ is complete, $\{A_n g\}$ converges to an element in $\mathcal{H}$ as $n \to \infty$, an element which may be defined as Ag. This proves the first part of the lemma: $A_n \to A$ on all $\mathcal{H}$ as $n \to \infty$.

2) To show that A is bounded with bound M, we notice that, since $A_n \to A$ as $n \to \infty$, corresponding to any $\varepsilon > 0$ and any $g \in \mathcal{H}$, N exists such that

$$\|Ag\| = \|(A - A_n + A_n)g\| \leqslant \|(A - A_n)g\| + \|A_n g\| \leqslant \varepsilon + M\|g\| , \qquad n > N .$$

Therefore, $\|A\| \leqslant M$.

Finally we have

Lemma 6. If for all $f \in \mathcal{D}$, where $\mathcal{D}$ is a linear manifold dense in $\mathcal{H}$, $\|Af\| \leqslant c\|f\|$, $\|A\| \leqslant M$, and c is independent of f, then $\|A\| \leqslant M \mathrm{in}\{M, c\}$.

Proof. It clearly suffices to assume that $M > 0$. By hypothesis corresponding to any $g \in \mathcal{H}$, $\{g_n\}$ exists with $g_n \in \mathcal{D}$ such that $g_n \to g$. Therefore for for any $\varepsilon > 0$, N exists such that $\|g - g_n\| < \varepsilon/(M + c)$, $n > N$. For such n we have the relations

$$\|Ag\| \leqslant \|A(g - g_n) + \|Ag_n\| \leqslant M \cdot \frac{\varepsilon}{M + c} + c\|g_n\| \leqslant$$

$$\leqslant \frac{M}{M + c}\varepsilon + c\|g - g_n\| + c\|g\| \leqslant \varepsilon + c\|g\|,$$

which shows that $\|A\| \leqslant c$. This result together with the assumption that $\|A\| \leqslant M$ proves the lemma.

On combining Lemmas 4, 5 and 6 we obtain Theorem 1. To see this we notice that $\{E_1 U_{t_n}\}$ satisfies the requirements of Lemma 5 if $t_n \to \pm\infty$ as $n \to \infty$, since 1), $\|E_1 U_{t_n}\| \leqslant |E_1| \cdot |U_{t_n}| \leqslant 1$, independently of n, and 2) $E_1 U_{t_n}$ satisfies the second requirement on account of the Corollary to Lemma 4. Therefore, $\{E_1 U_{t_n}\}$ converges on all $\mathcal{H}$ to an operator A such that $\|A\| \leqslant 1$. But on the dense domain of Lemma 4, A satisfies the requirements of Lemma 6 with $c = 0$. Therefore $E_1 U_{t_n} \to 0$ if $t_n \to \pm\infty$ as $n \to \infty$. This convergence

to zero is however independent of the particular sequence $\{t_n\}$. Therefore $E_\Delta U_t \to 0$ on all of $\mathcal{H}$ as $t \to \pm\infty$. (*q.e.d.*).

2. – Uniqueness of the free Hamiltonians of a scattering system.

The wave operator of a single-channel scattering system has been studied by many authors ([10]) and has been recently defined most succinctly and precisely by JAUCH ([3]). In a more recent paper ([2]) JAUCH has extended this formulation to the case of a multi-channel scattering system. Since in the present paper we are interested in such multi-channel scattering systems, we shall outline the definition of such systems and summarize the properties of the associated wave operators needed for the proof of Theorem 2.

A scattering system is characterized by a countable set of non-interacting fragments at $t = -\infty$ and by another such set at $t = +\infty$. During finite times the incoming fragments are in a finite region of space where they interact and yield a long time later ($t \to +\infty$) what is recognized experimentally as the second set. Each fragment at $t = \pm\infty$ is assumed to consist of a bound set of fundamental particles, and each fragment is characterized by a momentum operator $\boldsymbol{P}_i$, which is the sum of the momentum operators of the particles constituting it. The free or asymptotic motion of all the fragments is described by the self-adjoint Hamiltonian

$$H_\alpha = \frac{1}{2m_1} P_1^2 + \frac{1}{2m_2} P_2^2 + \dots \tag{8}$$

where α is a label specifying the channel energy, and the set $\{m_i\}$ consists of the masses of the fragments appearing in this channel. The self-adjoint Hamiltonian of the entire system is H and the interaction is defined by $H - H_\alpha$.

Let

$$U_t^{(\alpha)} = \exp[-iH_\alpha t]\,; \qquad V_t = \exp[-iHt]\,, \tag{9}$$

where $\{U_t^{(\alpha)}\}$ and $\{V_t\}$ are unitary groups which clearly satisfy the relation

$$[U_t^{(\alpha)}, U_{t'}^{(\beta)}]_- = 0\,, \qquad \text{for all } \alpha, \beta, t, t'\,. \tag{10}$$

It is further assumed or it can be proved from the assumptions ([2]) that if $\mathcal{H}$

([10]) C. MØLLER: *Kgl. Dan. Videnskab. Selskab, Mat.-fys. Medd.*, **23**, 1 (1945); B. LIPPMANN and J. SCHWINGER: *Phys. Rev.*, **79**, 469 (1950); M. GELL-MANN and M. L. GOLDBERGER: *Phys. Rev.*, **91**, 398 (1953).

is the underlying abstract Hilbert space, then (*A*) a set of bounded wave operators $\{\Omega_{\pm}^{(\alpha)}\}$ eists such that $\lim_{t\to\mp\infty} V_t^* U_t^{(\alpha)}$ exists on a closed domain $\mathcal{D}_\alpha$ and defines $\Omega_{\pm}^{(\alpha)}$ as a partial isometry (an isometry on $\mathcal{D}_\alpha$; 0 on the orthogonal complement, $\mathcal{D}_\alpha'$, of $\mathcal{D}_\alpha$); (*B*) the ranges of $\Omega_{\pm}^{(\alpha)}$ are $R_{\pm}^{(\alpha)}$ (and of course the $R_{\pm}^{(\alpha)}$ contain elements different from and in addition to the trivial null element) and the closed linear manifolds $R_\pm$ formed from the sets $\{R_{\pm}^{(\alpha)}\}$ and $\{R_{-}^{(\alpha)}\}$, respectively, are such that $R_+ = R_- = N$, where N is the sub-space of $\mathcal{H}$ associated with the continuous spectrum of H; (*C*) the number of channels is at most denumerably infinite. With these remarks in mind we may now state

THEOREM 2. Let $\{U_t^{(\alpha)}\}$ and $\{\overline{U}_t^{(\alpha)}\}$, $\alpha = 1, 2, \ldots$, be two sets of unitary groups such that

$$U_t^{(\alpha)} = \exp[-iH_\alpha t]\,; \qquad \overline{U}_t^{(\alpha)} = \exp[-i\overline{H}_\alpha t]\,, \tag{11}$$

where

$$H_\alpha = \sum_{j=1}^{n} a_j^{(\alpha)} \boldsymbol{P}_j^2\,; \qquad \overline{H}_\alpha = \sum_{j=1}^{n} \bar{a}_j^{(\alpha)} \boldsymbol{P}_j^2 \tag{12}$$

and $\{a_j^{(\alpha)}\}$, $\{\bar{a}_j^{(\alpha)}\}$ are sets of real numbers. Also let $\{V_t\}$ be a unitary group such that for each α $\lim_{t\to\mp\infty} V_t^* U_t^{(\alpha)}$ and $\lim_{t\to\mp\infty} V_t^* \overline{U}_t^{(\alpha)}$ exist on closed domains $\mathcal{D}_\alpha$ and $\overline{\mathcal{D}}_\alpha$, respectively, and define the partial isometries $\Omega_{\pm}^{(\alpha)}$ and $\overline{\Omega}_{\pm}^{(\alpha)}$, with ranges $R_{\pm}^{(\alpha)}$ and $\overline{R}_{\pm}^{(\alpha)}$, respectively. Finally assume that $R_+ = \overline{R}_+ = R_- = \overline{R}_- = N$. Then the sets $\{H_\alpha\}$ and $\{\overline{H}_\alpha\}$ are identical, except possibly for the order of their elements.

What this theorem states is that for any multi-channel scattering system as defined in (*A*)–(*C*) there is only one set of non-relativistic free Hamiltonians available to describe the free or asymptotic motion of the fragments. Expressed otherwise, the theorem states that *if* the total Hamiltonian can be split into two parts of which one part, H_α, is suitable for the definition of a scattering process, this splitting can be done in only one way. One should add to this formulation the provision that the H_α have the form of (12). However, it will be clear that the theorem can also be proved when relativistic expressions replace those in (12). Still it remains a moot point whether Theorem 2 is true if the form of H_α is unspecified. Relative to this more general problem Theorem 2 may be regarded as a lemma.

The proof of the theorem consists in showing (in Theorem 3) that if H_0 and $\overline{H}_0$ are of the form (12) such that $H_0 \neq \overline{H}_0$, and if in conjunction with a given unitary group $\{V_t\}$ they define wave operators $\Omega_{\pm}^{(0)}$ and $\overline{\Omega}_{\pm}^{(0)}$ as partial isometries, then the ranges $R_{\pm}^{(0)}$ and $\overline{R}_{\pm}^{(0)}$ of these wave operators are orthogonal to each other. Applied to the sets $\{H_\alpha\}$ and $\{\overline{H}_\alpha\}$ of Theorem 2, this result implies that if one α' exists such that $\overline{H}_{\alpha'} \neq H_\alpha$ for *all* α, $\overline{R}_{\pm}^{(\alpha')}$ is ortho-

gonal to all $R_{\pm}^{(\alpha)}$, that is, to N, and consequently also to $\overline{R}_{\pm}$ and finally to itself, since by assumption $\overline{R}_{\pm}^{(\alpha)} \subset R_{\pm}$. Therefore $R_{\pm}^{(\alpha)}$ contains no non-null element, contrary to the assumptions. Consequently $\{\overline{H}_\lambda\} \subset \{H_\alpha\}$. Similarly, one must conclude that $\{H_\alpha\} \subset \{\overline{H}_\lambda\}$. Therefore, $\{H_\alpha\} = \{\overline{H}_\lambda\}$. (*q.e.d.*). The proof of Theorem 2 will therefore be completed with the proof of

THEOREM 3. *Let* $\{U_t\}$ and $\{\overline{U}_t\}$ be two unitary groups defined for $-\infty < t < \infty$ such that

$$U_t = \exp[-iH_0 t]\,; \qquad \overline{U}_t = \exp[-i\overline{H}_0 t]\,, \tag{13}$$

where

$$H_0 = \sum_{i=1}^{n} a_i \boldsymbol{P}_i^2\,; \qquad \overline{H}_0 = \sum_{i=1}^{n} \overline{a}_i \boldsymbol{P}_i^2\,; \qquad a_i^2,\ \overline{a}_i^2 \geqslant 0. \tag{14}$$

Also, let $\{V_t\}$ be a unitary group $(-\infty < t < \infty)$ such that $\lim V_t^* U_t f$ and $\lim V_t^* \overline{U}_t \overline{f}$ when $t \to \mp\infty$ exist for arbitrary elements f and $\overline{f}$ in the closed domains $\mathcal{D}$ and $\overline{\mathcal{D}}$, respectively. Then if $H_0 \neq \overline{H}_0$, $R_- \perp \overline{R}_-$ and $R_+ \perp \overline{R}_+$, where R_-, $\overline{R}_-$, R_+, $\overline{R}_+$ are the ranges of the partial isometries Ω_-, $\overline{\Omega}_-$, Ω_+, $\overline{\Omega}_+$, respectively, which are defined by the three groups $\{U_t\}$, $\{\overline{U}_t\}$ and $\{V_t\}$.

The proof of this theorem will be carried out with the aid of three lemmas, the first of which is

Lemma 7. Let $\{U_t\}$ and $\{\overline{U}_t\}$ be two unitary groups defined for $-\infty < < t < \infty$ such that

$$U_t = \exp[-iH_0 t]\,; \qquad \overline{U}_t = \exp[-i\overline{H}_0 t]\,, \tag{15}$$

where

$$H_0 = \sum_{i=1}^{n} a_i \boldsymbol{P}_i^2\,; \qquad \overline{H}_0 = \sum_{i=1}^{n} \overline{a}_i \boldsymbol{P}_i^2\,; \qquad a_i^2,\ \overline{a}_i^2 \geqslant 0. \tag{16}$$

Then, if $H_0 \neq \overline{H}_0$,

$$\overline{U}_t^* U_t \rightharpoonup 0 \qquad \text{as} \qquad t \to \pm\infty\,. \tag{17}$$

Proof. Just as in the proof of Theorem 1 introduce as a realization of the underlying Hilbert space $\mathcal{H}$ the set $\mathcal{H}(\boldsymbol{k}_i)$ of L^2-functions defined on the entire $3n$-dimensional Euclidean space E_{3n}. $\mathcal{H}(\boldsymbol{k}_i)$ has properties similar to those espressed in (1) and (2), *i.e.*,

$$\omega(\boldsymbol{P}_1^2, \boldsymbol{P}_2^2, \ldots, \boldsymbol{P}_n^2)\varphi(\boldsymbol{k}_1, \boldsymbol{k}_2, \ldots, \boldsymbol{k}_n) = \omega(\boldsymbol{k}_1^2, \boldsymbol{k}_2^3, \ldots, \boldsymbol{k}_n^2)\varphi(\boldsymbol{k}_1, \boldsymbol{k}_2, \ldots, \boldsymbol{k}_n)\,, \tag{18}$$

where $\varphi(\boldsymbol{k}_1, \boldsymbol{k}_2, ..., \boldsymbol{k}_n) \equiv \varphi(\boldsymbol{k}_i)$ is an arbitrary element of $\mathcal{H}(\boldsymbol{k}_i)$ such that $\|\omega(\boldsymbol{P}_1^2)\varphi\| < \infty$.

Consider first two special elements f and g in $\mathcal{H}$ which are represented in $\mathcal{H}(\boldsymbol{k}_i)$ by the same step-function:

$$f(\boldsymbol{k}_i) = g(\boldsymbol{k}_i) = \begin{cases} 1, & \text{for } (\boldsymbol{k}_1, ..., \boldsymbol{k}_n) \in \Delta, \\ 0, & \text{otherwise}, \end{cases} \tag{19}$$

where Δ is a rectangular region in E_{3n} whose sides are parallel to the axes of E_{3n} and whose measure is finite. Then

$$I(t) \equiv (f, \overline{U}_t^* U_t g) = \int_{\Delta} \exp\left[-it \sum_{r=1}^{n} (a_r - \overline{a}_r)\boldsymbol{k}_r^2\right] \mathrm{d}\tau, \tag{20}$$

where $\mathrm{d}\tau$ is the $3n$-dimensional volume element in E_{3n}. $I(t)$ can be written as a 3-fold iterated integral, each of whose iterates is a one-dimensional integral having the form

$$J(t) = \int_{\sigma_1}^{\sigma_2} \exp\left[-itqk^2\right] \mathrm{d}k, \tag{21}$$

where $q = a_s - \overline{a}_s$ for some s and, if $\boldsymbol{k}_s^2 = \sum_{\alpha=1}^{3} k_{s\alpha}^2$, $k^2 = k_s^2$ for some s and α. Also, (σ_1, σ_2) is the interval over which k can vary in Δ. Since according to the assumptions $H_0 \neq \overline{H}_0$, at least one value of r exists for which $a_r - a_r \neq 0$. For such values of r

$$J(t) = \frac{1}{2\sqrt{q}} \int_{q\sigma_1^2}^{q\sigma_2^2} \frac{\exp\left[-i\lambda t\right]}{\sqrt{\lambda}} \mathrm{d}\lambda.$$

The range of integration may be extended from $-\infty$ to $+\infty$ if the integrand is replaced by the function $\psi(\lambda)$ defined in the following way:

$$\psi(\lambda) = \begin{cases} \dfrac{\exp\left[-i\lambda t\right]}{\sqrt{\lambda}}, & q\sigma_1^2 \leqslant \lambda \leqslant q\sigma_2^2, \\ 0, & \text{otherwise}. \end{cases}$$

Since $\psi(\lambda)$ is absolutely integrable over the whole line, $J(t) \to 0$ as $t \to \pm\infty$ according to the Riemann-Lebesgue lemma ([11]). For other values of r, *i.e.*,

([11]) E. C. Titchmarsh: *The Theory of Fourier Integrals* (Oxford, 1937).

those for which $a_r - \bar{a}_r = 0$, $q = 0$ and $J(t) | \leqslant \sigma_2 - \sigma_1$. Therefore $I(t) = (f, \bar{U}_t^* U_t g) \to 0$ as $t \to \pm\infty$.

Although this result has been obtained for the very special step-function (19), it is also true for arbitrary square-integrable step-functions, since the scalar product is bilinear in its two arguments. Since the set of square-integrable step-functions is dense in $\mathcal{H}$ ([12]), the lemma has been proved for elements in a domain $\mathcal{J}$, dense in $\mathcal{H}$, *i.e.*,

$$(22) \qquad (f, \bar{U}_t^* U_t g) \to 0 \qquad \text{as} \qquad t \to \pm\infty \qquad \text{for } f, g \in \mathcal{J}.$$

To complete the proof of this lemma we must show that (22) is true for elements in the closure $\bar{\mathcal{J}}$ of $\mathcal{J}$. This is most easily shown with the aid of the following variant of the Banach-Steinhaus Theorem.

Lemma 8. Let $\{T_n\}$ be a sequence of bounded operators such that

1) $\|T_n\| \leqslant M$ for all n

2) $(T_n p, q)$ converges as $n \to \infty$ for each p and q in a domain $\mathcal{J}$ which is dense in $\mathcal{H}$.

Then $T_n \to T$, where T is bounded and $\|T\| \leqslant M$. It will be noticed that this lemma differs from Lemma 5 only in the types of convergences considered. Neither lemma implies the other, however. Since the theorem is trivial if $M = 0$, we shall assume that $M > 0$.

Proof. Let f and g be any two non-null elements in $\mathcal{H}$. Then, since $\mathcal{J}$ is dense in $\mathcal{H}$, sequences $\{f_n\}$ and $\{g_n\}$ exist in $\mathcal{J}$ such that $f_n \to f$ and $g_n \to g$ as $n \to \infty$. Then

$$(23) \qquad \begin{aligned} |(T_n f, g) - (T_m f, g)| &= |((T_n - T_m) f, g)| \leqslant |((T_n - T_m)(f - f_r), g)| + \\ &+ |((T_n - T_m) f_r, g - g_s)| + |((T_n - T_m) f_r, g_s)| \leqslant \\ &\leqslant \|T_n - T_m\| \cdot \|f - f_r\| \cdot \|g\| + \|T_n - T_m\| \cdot \|f_r\| \cdot \|g - g_s\| + |((T_n - T_m) f_r, g_s)| \leqslant \\ &\leqslant 2M \cdot \|f - f_r\| \cdot \|g\| + 2M \|f_r\| \cdot \|g - g_s\| + |((T_n - T_m) f_r, g_s)| . \end{aligned}$$

Since $f_r \to f \neq 0$ and $g_s \to g \neq 0$ as $r, s \to \infty$, R and S exist such that $f_R \neq 0$ and

$$(24) \qquad \|f - f_R\| \leqslant \varepsilon / 6M \cdot \|g\| \; ; \qquad \|g - g_S\| \leqslant \varepsilon / 6M \cdot \|f_R\| .$$

According to assumption 2, N exists such that for fixed R and S and for all $n, m > N$

$$(25) \qquad |((T_n - T_m) f_R, g_S)| < \varepsilon / 3 .$$

([12]) E. J. McSHANE: *Integration* (Princeton, 1944), p. 230.

Setting $r = R$ and $s = S$ in (23) as determined in (24), we obtain

$$(26)\quad |(T_n f, g) - (T_m f, g)| \leqslant 2M \cdot \frac{\varepsilon}{6M \cdot \|g\|} \cdot \|g\| + \\ + 2M \cdot \|f_R\| \cdot \frac{\varepsilon}{6M \cdot \|f_R\|} + \frac{\varepsilon}{3} = \varepsilon\,, \qquad n, m > N.$$

If either or both the relations: $f = 0$, $g = 0$ hold, the left-hand side of (26) automatically vanishes. Therefore (26) is true for all elements in $\mathcal{H}$ and asserts that T exists such that $T_n \rightharpoonup T$ as $n \to \infty$.

The boundedness of T follows from the result just obtained and the relation

$$M^2 \|f_i\|^2 \geqslant |(T_n f, T_{n'} f)| \to (Tf, Tf) = \|Tf\|^2$$

as $n, n' \to \infty$, a relation which is correct for arbitrary elements of $\mathcal{H}$.

Corollary. If $\{T_n\}$ has the properties required in Lemma 8, and if in addition $(T_n p, q) \to 0$ as $n \to \infty$ for all p and q in $\mathcal{J}$, then $(T_n f, g) \to 0$ as $n \to \infty$ for all f and g in $\bar{\mathcal{J}} = \mathcal{H}$.

Proof. According to Lemma 8, $T_n \rightharpoonup T$ as $n \to \infty$ and T is bounded. Therefore, for arbitrary f and g, $(T_n f, g) \to (Tf, g)$ as $n \to \infty$. If f and $g \in \mathcal{J}$, $(Tf, g) = 0$ by assumption. If f and $g \in \bar{\mathcal{J}}$, they are each limits of sequences of elements in $\mathcal{J}$. And for elements of these sequences $(Tf_n, g_m) = 0$, where $f_n \to f$ and $g_n \to g$ as $n \to \infty$. Since T is bounded, (Tf_n, g_m) is continuous in f_n and g_m, and therefore [continuity of the scalar product] vanishes also when, $n, m \to \infty$. (*q.e.d.*).

The corollary just proved completes the proof of Lemma 7. This can be seen in the following way. In (22) it was shown that $(f, \bar{U}_t^* U_t g) \to 0$ as $t \to \pm\infty$ for f and $g \in \mathcal{J}$, where $\mathcal{J}$ is dense in $\mathcal{H}$. This result implies that for elements of $\mathcal{J}$, $(f, \bar{U}_{t_n}^* U_{t_n} g) \to 0$ as $n \to \infty$ if $t_n \to \pm\infty$ as $n \to \infty$. In addition, since U_t and $\bar{U}_t$ are unitary, $\|\bar{U}_t^* U_t\| \leqslant 1$. Therefore, since $\{\bar{U}_t^* U_t\}$ satisfies the requirements of the corollary, we conclude that $\bar{U}_{t_n}^* U_{t_n} \rightharpoonup 0$ as $n \to \infty$. Finally, since this result holds for arbitrary sequences $\{t_n\}$ such that $t_n \to \pm\infty$ as $n \to \infty$, $\bar{U}_t^* U_t \rightharpoonup 0$ as $t \to \pm\infty$; and this is just the conclusion of Lemma 7.

We now prove Lemma 9, the final lemma required for the proof of Theorem 3.

Lemma 9. Let $\{U_t\}$ and $\{\bar{U}_t\}$ satisfy the requirements of Lemma 7. In addition let $\{V_t\}$ be another unitary family defined for $-\infty < t < \infty$ with the property that for the two non-null elements f and $\bar{f}$ at least one of the

following two pairs of limits exists:

$$(27)_+ \qquad V_t^* U_t f \to f_- ; \qquad V_t^* \overline{U}_t \bar{f} \to \bar{f}_- \qquad \text{as } t \to +\infty,$$

$$(27)_- \qquad V_t^* U_t f \to f_- ; \qquad V_t^* \overline{U}_t \bar{f} \to \bar{f}_+ \qquad \text{as } t \to -\infty.$$

Then

$$(28)_+ \qquad (f_-, \bar{f}_-) = 0 ;$$

$$(28)_- \qquad (f_+, \bar{f}_+) = 0 .$$

Proof. Since the proof of $(28)_-$ is similar to that of $(28)_+$, we shall prove only $(28)_+$. According to $(27)_+$ and because the scalar product is continuous in its two arguments,

$$(f_-, \bar{f}_-) = \lim_{t,\, t' \to \infty} (V_t^* U_t f, V_{t'}^* \overline{U}_{t'} \bar{f}) .$$

Since the left-hand side is independent of the exact manner in which $t, t' \to \infty$, we may take $t = t'$. Then

$$(f_-, \bar{f}_-) = \lim_{t \to \infty} (V_t^* U_t f, V_t^* \overline{U}_t \bar{f}) = \lim_{t \to \infty} (\overline{U}_t^* U_t f, \bar{f}) = 0 .$$

The last equality follows from Lemma 7. (*q.e.d.*).

Theorem 3 is now effectively a trivial corollary of Lemma 9, since according to the lemma each element of $\mathcal{D}$ and of $\overline{\mathcal{D}}$ in Theorem 3 is mapped by the isometries into two corresponding elements, which are mutually orthogonal.

As stated earlier (see paragraph preceding the statement of Theorem 3), proof of Theorem 3 is equivalent with the proof of Theorem 2. With the proof of Theorem 3 we have therefore completed the proof of the main theorem.

* * *

It is a great pleasure to acknowledge my indebtedness to Professor J. M. JAUCH for introducing me to the problems of this paper, for emphasizing the practical power available in the formal theory of HILBERT space, and for many stimulating discussions. I also wish to thank him most warmly for the generous hospitality he made available to me at the State University of Iowa.

SUPPLEMENTO AL VOLUME XII, SERIE X
DEL NUOVO CIMENTO

N. 1, 1959
2° Trimestre

Gli organismi biologici come sistemi aperti stazionari nel modello teorico di L. von Bertalanffy (*).

R. MORCHIO

Istituto di Fisica Teorica dell'Università - Genova

(ricevuto il 4 Maggio 1959)

SOMMARIO. — 1. Introduzione. – 2. Le proposte del von Bertalanffy. – 3. Aspetti cinetici e termodinamici dei « sistemi aperti stazionari ». – 4. Problemi e conseguenze del modello proposto. – 5. Formalizzabilità e accordo con l'esperienza del modello. – 6. Operatività del modello. – 7. Conclusione.

1. – Introduzione.

Il lavoro del biologo che affronta i tradizionali problemi della biologia secondo le tradizionali modalità, appare, a prima vista, metodologicamente diverso non soltanto dal tipo di lavoro del fisico, ma anche del biofisico o del biochimico, che, nei capitoli più recenti delle scienze biologiche, sempre più tendono ad un ampio uso della formalizzazione matematica.

Se però l'oggetto ultimo della ricerca dei biologi, dei biochimici e dei biofisici è fondamentalmente lo stesso, perchè questa unità di ricerca non sia limitata a contenuti più intuitivi che teorici, è necessario che proprio sul piano metodologico si arrivi ad un sostanziale chiarimento nel senso del superamento di ogni residuo dogmatismo sulla non applicabilità di metodi rigorosi e quantitativi [1] alla ricerca biologica a tutti i suoi livelli, che è poi quanto riaffermare in altri termini la omogeneità sostanziale della ricerca sulla natura.

(*) Lavoro fatto con un contributo del C.N.R.

(1) Nota giustamente il WOODGER (*Biology and Language*. Cambridge, (1952)) che veramente il problema non è tanto quello di rendere la biologia più quantitativa, quanto di trasformare i metodi « intensionali » in « estensionali ». Essendo però la matematica essenzialmente « estensionale » in prima approssimazione le due affermazioni non differiscono di molto.

L'affermazione dell'applicabilità dei metodi quantitativi alla ricerca biologica implica la effettiva applicazione concreta di tali metodi ai tradizionali problemi biologici (essa è già avvenuta in qualche singolo caso e per il resto sarà frutto della ricerca); ma può essere sostenuta, in linea teorica, anche mediante la dimostrata proponibilità di modelli teorici dell'organismo, che, suscettibili di formulazione matematica, siano in accordo con l'esperienza. In questo senso vale la pena esaminare se esista qualche modello rispondente alle condizioni accennate: che non sia cioè di solo interesse filosofico [2] ma che possa costituire piuttosto una ipotesi di lavoro.

2. – Le proposte del von Bertalanffy.

Ludwing von Bertalanffy, il noto biologo viennese che vive attualmente negli Stati Uniti, ha proposto già da molti anni di considerare l'organismo biologico come un « sistema », complesso fin che si vuole, rispondente a sue particolari leggi di comportamento, ma aggredibile e studiabile con metodi di tipo fisicalistico. Naturalmente la parola « sistema » non bastava. Bisognava rispondere a problemi e difficoltà ben note, lasciando spazio per le soluzioni originali che i problemi originali dell'organismo biologico richiedono. Per questo il von Bertalanffy fin dal 1932 [3] avanzò la concezione dell'organismo come un « sistema aperto », un sistema cioè in cui lo scambio di materia con l'esterno modificava completamente lo schema del « sistema chiuso » descritto dalla fisica ed in particolare dalla termodinamica ordinaria.

Le difficoltà in ogni modo erano notevoli. Da un lato la biologia ha spesso usato espressioni come « sistema », « organismo », « interazione », « il tutto è più che la somma delle sue parti », « equilibri biologici », e simili, ma sempre, come nota lo stesso von Bertalanffy [4], in modo vago e talora mistico; dall'altro la stessa descrizione dei « sistemi chiusi » della fisica era assolutamente inapplicabile agli organismi viventi e non poteva dare molti aiuti. Scrive il Prigogine, a proposito della termodinamica: « Una seria limitazione della termodinamica classica come strumento generale per una descrizione macroscopica dei processi fisico-chimici sta nel fatto che il suo metodo è basato sui concetti di « processi reversibili » e sugli stati di vero equilibrio » [5]. Non a caso

[2] Anche il lavoro del Mainx (*Foundations of Biology*, in: *Intern. Encycl. of Unified Science*, vol. 1, parte 2ª, (Chicago, 1955), p. 567), pur interessante per molti aspetti, non soddisfa questo punto di vista.

[3] L. von Bertalanffy: *Theoretische Biologie*, Band II, I ed. (Berlin, 1932-42) e *Naturwiss.*, **28**, 521 (1940).

[4] L. von Bertalanffy: *Brit. Journ. Phil. Sci.*, **2**, 134 (1950).

[5] I. Prigogine: *Introduction to Thermodynamics of Irreversible Processes* (Springfield, 1955), p. v.

proprio su questo fatto nacque l'equivoco di considerare la termodinamica ordinaria, a causa della sua inapplicabilità agli organismi viventi, quale prova della differenza sostanziale e incolmabile tra vita e non vita. Il Lecomte du Noüy, in particolare, diffuse tra i biologi questo argomento che acquistò in molti ambienti un notevole prestigio, anche se immeritato.

Per tutti questi motivi il VON BERTALANFFY da circa vent'anni ha sottolineato la necessità di generalizzare alcuni dei concetti della fisica e della chimica, perchè potessero essere estesi anche ai « sistemi aperti ». Lo stesso VON BERTALANFFY [6] e il BURTON [7] cominciarono a sviluppare qualche principio della cinetica dei « sistemi aperti »: mentre i processi irreversibili e i « sistemi aperti » stessi sono stati studiati termodinamicamente in particolare dalla scuola belga: DONDER, ONSAGER, PRIGOGINE e altri [8].

Diveniva così possibile al VON BERTALANFFY raccogliere tutti questi elementi e delineare un primo schema di un modello d'interpretazione dell'organismo vivente come « sistema aperto ».

Per la verità questa visione dell'organismo fu ampiamente generalizzata dal VON BERTALANFFY. La sua conclusione fu che ciascuna scienza studiando dei « sistemi » e presentando relazioni matematiche, leggi, principi formalmente equivalenti, sarebbe possibile formulare alcune relazioni generalissime, le quali, riferendosi ad un sistema in generale, potrebbero valere per tutte le scienze, rendendo finalmente possibile, su un piano formale, un'effettiva unità delle scienze stesse. Tali concetti costitutiscono il fondamento di quella che il VON BERTALANFFY chiamò « General System Theory », strumento, nel pensiero dell'autore, di controllo e di spinta a trasferire principi da un campo all'altro. La estrema generalità della « General System Theory » [9] la pone, in questa sua forma, al di fuori dei limiti e dell'interesse di questo lavoro.

Qui non si vuole, del resto nè esporre dettagliatamente il modello dell'organismo come « sistema aperto », nè analizzarne la validità in modo esauriente. Si vuole solo esaminare se esso è proponibile. Per questo è sufficiente un breve « survey » della teoria e delle eventuali conferme. Sarà l'affermazione di proponibilità del modello che potrà aprire, in un secondo tempo, un discorso critico più impegnato.

[6] L. VON BERTALANFFY: *Naturwiss.*, **28**, 521 (1940), già citato.

[7] A. C. BURTON: *Journ. Cell. Comp. Physiol.*, **14**, 327 (1939).

[8] Vedi in particolare: I. PRIGOGINE: *Étude thermodynamique des phénomènes irréversibles* (Parigi, 1947) e *Introduction to Thermodynamics of Irreversible Processes* (già citato); S. R. DE GROOT: *Thermodynamics of Irreversible Processes* (New York, 1951); I. PRIGOGINE e J. M. WIAME: *Experientia*, **2**, 450 (1946).

[9] Il VON BERTALANFFY, citando il filosofo N. HARTMANN, parla della G.S.T. addirittura come di uno strumento per rimpiazzare la teoria delle categorie con un sistema esatto di leggi logico-matematiche.

Ma che significa chiedersi se questo modello è proponibile? Nello spirito di quanto è stato detto si possono delineare i seguenti criteri: *a*) che il « modello » sia formalizzabile, sia suscettibile cioè di una formulazione matematica, *b*) che non sia, ovviamente, in disaccordo con l'esperienza, e *c*) che sia operativo, che permetta cioè di trarre dalla sua formulazione generale conseguenze passibili di verifica sperimentale.

I punti essenziali del modello, che verranno qui brevemente proposti, saranno valutati appunto in base a questi criteri.

3. – Aspetti cinetici e termodinamici dei « sistemi aperti stazionari ».

Il von Bertalanffy intende in generale per « sistema »: « Un complesso di elementi $p_1, p_2, \ldots, p_n$ interagenti. Interazione significa che gli elementi stanno in una certa relazione R, cosicchè il loro comportamento in R è diverso dal loro comportamento in un'altra relazione R'. Se il comportamento in R, ed in R' non è differente, non vi è interazione e gli elementi si comportano indipendentemente rispetto alle relazioni R e R' » [10].

Un tale sistema può biologicamente riferirsi, a livelli diversi, ad una popolazione (gli individui ne costituiscono gli elementi), ad un organismo solo (gli organi, o le cellule ne costituiranno gli elementi), ad una cellula, ecc. Tali sistemi possono essere descritti da sistemi di equazioni ben note. Indicando la misura di qualche aspetto quantitativo degli elementi p di cui sopra con $Q_1, Q_2, Q_3, \ldots, Q_n$, la sua variazione può essere definita appunto da un sistema di equazioni differenziali simultanee che per un numero finito di elementi, nel caso più semplice, avrà la forma:

$$(1)\qquad \begin{cases} \dfrac{\mathrm{d}Q_1}{\mathrm{d}t} = f_1(Q_1, Q_2, \ldots, Q_n), \\ \dfrac{\mathrm{d}Q_2}{\mathrm{d}t} = f_2(Q_1, Q_2, \ldots, Q_n), \\ \quad \ldots \quad \ldots \quad \ldots \quad \ldots \\ \dfrac{\mathrm{d}Q_n}{\mathrm{d}t} = f_n(Q_1, Q_2, \ldots, Q_n). \end{cases}$$

Tale sistema esprime, secondo von Bertalanffy, un principio generale di cinetica valido per qualunque sistema; ovvio che, a livelli diversi, sarà diversa l'applicazione, come diverso sarà il significato dei parametri e delle variabili.

Chiaro è il significato di queste equazioni. È pertanto sufficiente qui ricor-

[10] L. von Bertalanffy: *Brit. Journ. Phil. Sci.*, **2**, 134 (1950), già citato.

darne le conseguenze solo per alcuni casi particolarmente semplici ma non privi di importanza per la biologia, come si vedrà successivamente.

Se il sistema studiato è composto di elementi di una sola specie. l'(1) diviene semplicemente:

$$\frac{dQ}{dt} = f(Q)$$

che, supposta la funzione analitica, sviluppata in serie di Taylor nell'intorno di $Q = 0$, diviene:

$$\frac{dQ}{dt} = a_1 Q + a_{11} Q^2 + \dots . \tag{2}$$

Se della (2) si considera soltanto il primo termine del membro di destra, se cioè l'accrescimento (o il decremento) numerico di Q dipende esclusivamente in qualche maniera dallo stesso valore che assume Q in ogni singolo momento, si ha il caso più semplice:

$$\frac{dQ}{dt} = a_1 Q ,$$

cioè integrando:

$$Q = Q_0 \exp [a_1 t] ,$$

indicante un incremento o un decremento (a_1 positivo o negativo) esponenziale.

Considerando invece i primi due termini del membro di destra della (2), nel caso che l'accrescimento numerico (o il decremento) di Q sia funzione non soltanto del valore dello stesso Q, ma anche di un altro fattore, si ha:

$$\frac{dQ}{dt} = a_1 Q + a_{11} Q^2 ,$$

che, integrata, dà la soluzione:

$$Q = \frac{a_1 C \exp [a_1 t]}{1 - a_{11} C \exp [a_1 t]} ,$$

descrivente un accrescimento indicato da una curva logistica.

Nel caso di un sistema consistente di due specie di elementi, il sistema si svolgerà secondo linee diverse a seconda dei casi. Sarà rappresentato dalle due equazioni:

$$\begin{cases} \dfrac{dQ_1}{dt} = f_1(Q_1, Q_2) , \\ \dfrac{dQ_2}{dt} = f_2(Q_1, Q_2) , \end{cases}$$

581

che, svolte in serie di Taylor e risolte, in base al valore assunto caso per caso dai parametri, daranno tre possibilità: 1) descrizione di un sistema che tende verso uno stato stabilmente stazionario (si vedrà poi quale significato esatto assuma la parola « stazionario »), 2) descrizione, mediante termini periodici, di un sistema che oscilli attorno a valori stazionari, 3) descrizione di un sistema non in stato stazionario, nè oscillatorio.

Gli aspetti termodinamici dei « sistemi aperti » (ne abbiamo visto quelli cinetici) devono essere esaminati tenendo conto, come s'è detto della recente termodinamica dei processi irreversibili. Di essa basta ricordare qui alcuni punti, rimandando ai testi indicati nella nota ([8]) per una maggior documentazione.

La termodinamica ordinaria ha a che fare con sistemi chiusi, sistemi per i quali non si ha scambio di materia con l'ambiente esterno, ma solo di energia. Per essi vale la relazione di Carnot-Clausius:

$$\mathrm{d}S = \mathrm{d}_e S + \mathrm{d}_i S = \frac{\mathrm{d}Q}{T} + \mathrm{d}_i S\,,$$

in cui S indica l'entropia, gli indici e ed i si riferiscono rispettivamente all'entropia scambiata dal sistema con l'esterno (e) e a quella prodotta all'interno (i) del sistema dai processi irreversibili, $\mathrm{d}Q$ è il calore fornito al sistema dall'ambiente e T la temperatura (assoluta). $\mathrm{d}_i S$ può essere soltanto zero o positivo. La termodinamica ordinaria considera il caso in cui $\mathrm{d}_i S = 0$ che dà la seconda legge per i processi reversibili ($\mathrm{d}S = \mathrm{d}Q/T$).

Se abbiamo $\mathrm{d}_i S > 0$, cioè se $\mathrm{d}S > \mathrm{d}Q/T$, che si riferisce evidentemente ad un processo irreversibile, la termodinamica ordinaria descrive quest'ultimo come un cambiamento infinitamente lento, da uno stato di equilibrio ad un altro. Questo schema, per essere utilizzabile biologicamente, deve evidentemente esser reso più complesso.

Già qualche sistema chiuso può ampliarlo: il secondo principio richiede che in un sistema chiuso la produzione di entropia risultante da eventuali reazioni simultanee sia positiva. Può però accadere che, indicando con A l'affinità chimica e con v la velocità di reazione, un sistema subisca due reazioni simultanee tali che $A_1 v_1 < 0$ e $A_2 v_2 > 0$, purchè $A_1 v_1 + A_2 v_2 > 0$. Tali reazioni sono dette *reazioni accoppiate*, ed è così possibile che una delle reazioni progredisca in direzione contraria a quella prescritta dalla sua affinità. Così nella termodiffusione, ad esempio, si ha una diffusione di materia che avviene contro il suo gradiente di concentrazione, producendo entropia negativa, compensata peraltro dalla produzione di entropia positiva dovuta al flusso di calore. Tale descrizione ha già un interesse biologico perchè negli organismi viventi sono estremamente importanti le reazioni accoppiate.

Ma ancor più interessante sarà questa descrizione se, prendendo in consi-

derazione i sistemi aperti, la relazione di Carnot-Clausius, in $d_e S$, verrà completata con un termine ulteriore rappresentante la fornitura di entropia da parte dell'ambiente attraverso il trasporto di materia. Ora se l'entropia assunta dal sistema aperto in relazione ad una assunzione di materia è negativa e pareggia la produzione all'interno del sistema, si ha una situazione che non è di equilibrio nel senso della termodinamica dei processi reversibili perchè il sistema subisce delle trasformazioni, ma queste trasformazioni sono tali per cui il loro risultato complessivo è di mantenere il sistema in una situazione costante. La termodinamica dei processi irreversibili chiama questo stato *stazionario*.

Un esempio di stato stazionario può esser dato da un sistema che riceve in continuazione un componente M dell'ambiente esterno e lo trasforma continuamente, attraverso un certo numero di composti intermedi, in un prodotto finale F che è restituito in continuazione all'ambiente esterno, mentre le concentrazioni dei componenti intermedi non variano col tempo. La condizione perchè si abbia appunto lo stato stazionario stà nel fatto che la produzione di entropia abbia il suo valore minimo compatibile con le condizioni stabilite caso per caso dalla concreta situazione e che, nell'esempio testè presentato, potrebbero essere le concentrazioni dei prodotti iniziali e finali M e F nell'ambiente esterno cui il sistema deve adeguarsi.

I problemi relativi agli stati stazionari, per il loro interesse, sono stati discussi non solo da fisici, ma anche da biologi ([11]). La termodinamica dei processi irreversibili ha portato infatti a conseguenze di non poco interesse biologico, dimostrando che, operando sempre i processi irreversibili interni in modo da abbassare il valore della produzione di entropia per unità di tempo, quando un sistema è in uno stato di minima produzione di entropia, allora non può lasciare questo stato mediante un cambiamento irreversibile spontaneo. Se per qualche fluttuazione devia leggermente da questo stato, avranno luogo cambiamenti interni che riporteranno indietro il sistema al suo stato iniziale, il quale può considerarsi perciò uno stato stabile come pure stabili possono considerarsi le trasformazioni che appunto in tale stato intervengono.

In generale si può dire dunque, più esattamente, che in un sistema caratterizzato da n forze indipendenti $X_1, X_2, \ldots, X_n$, lo stato stazionario consiste nella presenza di un certo numero di costrizioni, fissate da un certo numero delle n forze $X_1, X_2, \ldots, X_k$ (k è uno dei numeri $0, 1, 2, \ldots, n$) tenute a valori costanti (stato stazionario di ordine k). In un tale sistema, con una produzione minima di entropia σ, si vanificano i flussi J_j con gli indici $j = k+1, k+2, \ldots, n$, cioè: $J_{k+1} = J_{k+2} = \ldots = J_n = 0$, dove ciascun J_j è il flusso legato alla relativa forza X_j. Questo teorema, che fu formulato per primo da Pri-

([11]) Tra i biologi sono da ricordare: A. J. Lotka: *Proc. Nat. Acad. Sci.*, **8**, 146 (1922); *Bull. Math. Biophys.*, **10**, 103 (1948) e J. Z. Hearon: *Bull. Math. Biophys.*, **12**, 57 (1950).

gogine per il caso di un singolo parametro fissato, è una generalizzazione del principio di moderazione di Le Chatelier.

È ovvio che in uno stato stazionario una produzione di entropia dev'essere compensata da un flusso negativo di entropia in modo che la variazione totale di entropia nel tempo sia zero:

$$\frac{\mathrm{d}S}{\mathrm{d}t} = \frac{\mathrm{d}_i S}{\mathrm{d}t} + \frac{\mathrm{d}_e S}{\mathrm{d}t} = 0 ,$$

Ma:

$$\frac{\mathrm{d}_i S}{\mathrm{d}t} > 0 ,$$

per cui:

$$\frac{\mathrm{d}_e S}{\mathrm{d}t} < 0 .$$

L'entropia della materia che entra in un sistema in stato stazionario deve perciò essere più piccola dell'entropia della materia che il sistema restituisce al mondo esterno. Si ha perciò una degradazione della materia che entra nel sistema, degradazione che mantiene lo stato stazionario, contribuendo all'aumento dell'entropia esterna. Durante l'evoluzione verso lo stato stazionario la produzione di entropia e spesso la stessa entropia decrescono prendendo il valore più basso compatibile con le costrizioni esterne quando lo stato stazionario è raggiunto.

Questo tipo di descrizione è particolarmente efficace per il comportamento globale degli organismi viventi a differenza della descrizione dei sistemi della termodinamica ordinaria [12]. L'evoluzione degli organismi fino allo stato adulto può vedersi come l'evoluzione del sistema verso lo stato stazionario, avente luogo sotto un certo numero di costrizioni determinate dal mondo esterno e questo comporta appunto una diminuzione nella produzione di entropia. Il fatto che l'organizzazione di un simile sistema durante l'accrescimento aumenti, spiega la diminuzione di entropia, mentre la sua stabilità è chiarita dalla stabilità dei processi stazionari e la sua capacità a riequilibrarsi, dopo modificazioni non eccessive subite, è spiegata in base al principio di Le Chatelier generalizzato. Il continuo dinamismo dei processi metabolici in cui quelli anabolici pareggiano quelli catabolici ed il continuo rinnovarsi della materia componente gli organismi, coesistenti con la permanenza della forma e delle strutture così autoregolantisi, trovano dunque una chiarificazione nella dinamica degli stati stazionari.

[12] Si veda in particolare lo scritto di Prigogine e Wiame in *Experientia* (1946) già citato. Anche: L. von Bertalanffy: *Science*, **3**, 23 (1950).

4. – Problemi e conseguenze del modello proposto.

Il modello proposto consiste perciò nel considerare un organismo vivente come un sistema aperto in uno stato stazionario e nel descrivere tale sistema cineticamente con le equazioni differenziali indicate all'inizio della Sezione 3 e termodinamicamente con le equazioni della termodinamica dei processi irreversibili.

Tale descrizione è sostanzialmente omogenea non sussistendo contraddizioni tra i due tipi di descrizione. Non è qui il caso, naturalmente, di esaminare se tale descrizione implichi o meno l'assioma di causalità come qualcuno ha fatto ([13]), ma è utile sottolineare come questa sia una descrizione globale del comportamento di un organismo; o anche di sotto-sistemi di un organismo, ma sempre nei loro comportamenti globali. Questo modo di vedere un organismo, a nostro avviso, apre proprio il problema di come questo modello possa eventualmente accordarsi con la considerazione di singoli eventi microscopici, notevolmente importanti in un vivente.

Si può anche porre il quesito di quanto grande sia il contributo del VON BERTALANFFY nell'elaborazione di questo modello. La termodinamica dei processi irreversibili è stata sviluppata da un gruppo di fisici. Le equazioni differenziali cinetiche non si devono completamente a lui; del resto la legge dell'azione di massa poteva già dare una indicazione. Il VON BERTALANFFY però, come è stato detto, già da tempo aveva intuito che un organismo poteva considerarsi come sistema aperto, prima che la termodinamica dei processi irreversibili fornisse precise indicazioni in proposito. Inoltre, e non è poco, il Nostro raccolse tutti questi elementi e contribuì notevolmente a chiarirne l'interesse biologico ed a farli conoscere ai biologi. Infine utilizzò queste idee proprio all'interno delle classiche concezioni della biologia tradizionale. Su quest'ultimo punto, sia pure brevemente, vale la pena spender qualche parola ([14]).

L'unitarietà del comportamento di un organismo vivente assume, ad esempio, per VON BERTALANFFY un significato ben preciso cineticamente e termodinamicamente. Certe espressioni generiche come: « una modificazione in una parte influenza il tutto », ecc., si traducono da un lato nell'affermazione termodinamica secondo cui una qualsiasi modificazione nel sistema determina complessivamente nel sistema stesso tali modificazioni da ristabilire l'equilibrio, dall'altro lato nella constatazione che dall'equazione, sviluppata in serie di Taylor:

$$\frac{dQ_i}{dt} = a_{i1}Q_1 + a_{i2}Q_2 + \ldots + a_{in}Q_n + a_{i11}Q_1^2 + \ldots ,$$

([13]) W. YOURGRAU: *Scientia*, **11**, 307 (1952).

([14]) Indicazioni maggiori possono vedersi in: L. VON BERTLANFFY: *Brit. Journ. Phil. Sci.*, 2, 134 (1950), già citato; *Human Biology*, **23**, 302 (1951); *Journ. Personality*, **20**, 24 (1951); *Main Currents in Modern Thought*, **4**, 75 (1955).

risulta che ogni mutamento di una quantità Q_1, è funzione polinomiale di tutti gli elementi Q_1, Q_2..., Q_n. A sua volta un mutamento di uno dei Q_i causa una variazione in tutti gli altri Q_i e quindi nel sistema nel suo complesso. Di particolare interesse è l'applicazione che il VON BERTALANFFY fa di questi concetti a quel tipico comportamento biologico che il Nostro ha definito « equifinalità ». In un sistema fisico ordinario lo stato finale è determinato dalle condizioni iniziali; in un organismo vivente invece lo stato finale può essere raggiunto partendo da condizioni iniziali diverse e attraverso vie diverse. Il famoso esperimento su un germe di riccio di mare in base al quale lo stesso stato adulto è raggiunto partendo dal germe normale completo, quanto da una metà, dopo una divisione sperimentale, fu, come è noto, la base per la formulazione delle teorie vitalistiche di Driesch. Il VON BERTALANFFY vede nell'autoregolazione globale, risultante dalle caratteristiche sopra descritte del sistema aperto in stato stazionario, la spiegazione di questo fondamentale aspetto degli organismi. L'accrescimento è infatti il risultato del metabolismo, inteso come una contro-azione tra anabolismo e catabolismo dei materiali di costruzione, e il rapporto relativo è regolabile in termini di autoregolazione dei sistemi considerati.

Il modello del sistema aperto in stato stazionario conduce, come s'è visto, all'ammissione di meccanismi autoregolativi impliciti nel sistema. Ciò si accorda con le caratteristiche presentate dagli organismi. Ma un modello di autocontrollo valido per gli organismi è presentato anche dalla cibernetica. Si tratta dei meccanismi a « feed-back ». Il VON BERTALANFFY non vede alcuna contraddizione tra i due modelli. Il problema è trattato ripetutamente ([15]) e il Nostro conclude che certamente vi sono molti fenomeni biologici inquadrabili nello schema del « feed-back », dai fenomeni di omeostasi (la termoregolazione, ad esempio, negli animali a sangue caldo), ai meccanismi del sistema nervoso. Questi, per potersi verificare, presuppongono però arrangiamenti strutturali. Negli organismi esistono invece molte regolazioni di natura diversa. Esse non si basano su strutture preesistenti in qualche modo, ma si fondano sull'interazione dinamica dei processi. Sono esse, come dice il VON BERTALANFFY, *regolazioni primarie* e sono spiegabili in termini di stati stazionari o di tendenza allo stato stazionario in sistemi aperti. Le *regolazioni secondarie*, del tipo « feed-back », sovrimporrebbero successivamente, sia nel senso dello sviluppo embrionale, quanto in quello evolutivo, alla regolazione dinamica dei componenti, arrangiamenti fissi e condizioni di costrizione, rendendo il sistema e le sue parti più efficienti.

In *Problems of Life* ([16]) VON BERTALANFFY scrive: « A living organism is a

([15]) Vedi in particolare: L. VON BERTALANFFY: *Human Biology*, **23**, 346 (1951).

([16]) L. VON BERTALANFFY: *Problems of Life. An Evaluation of Modern Biological Thought* (Londra, 1952). Ed. originale: *Das biologische Weltbild* (Berna, 1949).

hierarchy of open systems maintaining itself in a steady state due to its inherent system conditions ». Anche altrove ritorna su quello che chiama: « ordine gerarchico » secondo cui, in molti sistemi, i componenti sono essi stessi sistemi di un ulteriore ordine più basso, significando con ciò che ogni Q_i è esso stesso un sistema di elementi R_i definibile pure da un sistema di equazioni cinetiche quali quelle già usate. A parte l'espressione « ordine gerarchico » che, se non definita rigorosamente, rischia di riportare il discorso su di un piano qualitativo, questo concetto di « sistemi di sistemi » non è stato approfondito dal VON BERTALANFFY quanto meriterebbe, e non consente pertanto qui un lungo esame.

5. – Formalizzabilità e accordo con l'esperienza del modello.

Occorre adesso tentare di rispondere alla domanda posta all'inizio. È questo modello proponibile? Si può confrontare con le tre condizioni poste. Si risponde subito facilmente alla prima. Il modello è certamente suscettibile di una formulazione matematica. Sia sotto il profilo cinetico, sia sotto quello termodinamico è utilizzabile il formalismo matematico per corrispondere al quale, in fondo, il modello è stato formulato.

Anche la seconda condizione non richiede un lunghissimo discorso. Il modello non è contraddetto da alcun fatto sperimentale specifico. Le stesse idee fondamentali su cui si basa oggi la biologia non sono in contrasto con esso. La teoria cellulare vi si inquadra perfettamente; anzi, si potrebbe dire, è esplicitamente richiesta dalla concezione di sistema composto di elementi, descritto dalle equazioni cinetiche scritte sopra. L'accrescimento individuale e l'evoluzione, cioè l'aumento dell'ordine è ammesso dalla termodinamica dei processi irreversibili, purchè siano rispettate le condizioni cui si è accennato sopra. Un discorso un po' diverso si potrebbe fare per la genetica e i problemi relativi. Il VON BERTALANFFY pur avendo studiato geni e cromosomi alla luce del modello considerato (vedi oltre), non ci pare abbia affrontato sistematicamente il problema di come dei sistemi aperti (geni e cromosomi) possono controllare il sistema più ampio (organismo) di cui sono elementi. È vero che (vedi sopra) secondo VON BERTALANFFY ogni mutamento di qualche quantità Q_i causa un cambiamento in tutti gli altri elementi e quindi nel sistema nel suo complesso. Si può anche ammettere che il Q_i, rappresentante un'unità genetica, appaia nell'equazione con parametri tali, rispetto agli altri elementi, che sue eventuali modificazioni abbiano un'influenza particolarmente ampia su altri elementi, o su alcuni tra questi. Però una simile formulazione è eccessivamente generica e insufficiente di fronte all'importanza ed alla regolarità che gli eventi genetici hanno negli organismi viventi. In ogni modo si tratta qui di una insufficienza piuttosto che di una smentita.

Non si può d'altra parte non sottolineare le non trascurabili conferme e concordanze che su singoli punti il modello presentato ha trovato. Ricorderemo qui alcuni dati.

Per quanto riguarda gli aspetti termodinamici già la termodinamica ordinaria permetteva di sviluppare, in casi particolari, concetti che saranno pienamente compresi alla luce della termodinamica nuova e del modello del « sistema aperto ». La termodinamica classica mostrava che talora è possibile trascurare, nelle equazioni relative, il termine della variazione dell'entropia. In questi casi la produzione di entropia dovuta ad una reazione chimica è semplicemente proporzionale al calore di reazione. In tale approssimazione la produzione di entropia di un organismo vivente può essere misurata dal suo metabolismo in base alle registrazioni calorimetriche. Una reazione accoppiata del tipo $Av < 0$ sarà perciò equivalente ad una reazione endotermica e la sua presenza si manifesterà mediante la diminuzione del calore totale che il sistema dà al mondo esterno. Molti autori hanno lavorato a questo problema, particolarmente nella fase embrionale, quando, per i processi di organizzazione, reazioni accoppiate devono giocare un ruolo particolarmente importante ([17]).

Gli aspetti biologici di un altro punto della termodinamica sono stati sviluppati dal LIPMAN ([18]). Per reazioni vicine all'equilibrio si ha $|A/RT| \ll 1$ (R è la costante dei gas). Se l'affinità di una data reazione chimica è grande e questa relazione non è soddisfatta, la reazione può spesso esser divisa in un certo numero di reazioni elementari, aventi ciascuna affinità sufficientemente piccola per giustificare l'applicazione delle leggi della termodinamica (le così dette leggi fenomenologiche lineari). Lo sviluppo di questi casi interessa il biologo proprio perchè costui ha spesso a che fare con reazioni chimiche che usualmente hanno luogo in molti passi, ciascuno dei quali è vicino alla reversibilità.

Ricordiamo infine che abbastanza recentemente HEARON ([19]) ha messo in evidenza l'importanza, per i processi biologici, dell'accoppiamento tra diffusione e reazioni chimiche, studiato dalla termodinamica e in base a cui si può capire, in una cellula vivente, una distribuzione stazionaria (non di equilibrio) di materia che differisca da quella ambientale in base ai rapporti metabolici della cellula stessa.

Anche la validità delle equazioni cinetiche usate dal VON BERTALANFFY è stata verificata in biologia. In capitoli diversi, studiosi diversi, per risolvere problemi diversi, hanno usato questo particolare tipo di equazioni differenziali.

Un caso notissimo e di particolare importanza è costituito dagli studi del

([17]) Vedi in particolare: J. NEEDHAM: *Chem. Embryol.*, 3 vols. (Cambridge, 1931).
([18]) F. LIPMAN: *Adv. Enzym.*, **1**, 99 (1941).
([19]) J. Z. HEARON: *Bull. Math. Biophys.*, **12**, 135 (1950).

Lotka, ma particolarmente del Volterra ([20]), sulla dinamica delle popolazioni. Non è evidentemente qui possibile soffermarsi su questi problemi, ma è sufficiente ricordare che l'accrescimento teorico di una specie, supposta isolata e non impedita da alcuna limitazione, si sviluppa proprio secondo l'equazione $dQ/dt = a_1 Q$ che, come s'è visto, integrata dà $Q = Q_0 \exp[a_1 t]$, indicante, per Q = numero degli individui al tempo t e Q_0 = numero iniziale degli individui, un aumento secondo una curva esponenziale. Con esponente negativo descrive invece il decremento di una popolazione in cui la mortalità sia maggiore della natalità, ma si applica anche all'uccisione di batteri da parte di disinfettanti o alla perdita di sostanza corporea in un animale che digiuni. In pratica però una popolazione, anche supposta isolata, oltre certi limiti, trova sempre ostacoli al suo sviluppo che pertanto va gradatamente diminuendo, sì da descrivere piuttosto che una curva esponenziale, una curva logistica. Anche questa curva è descritta analiticamente da un'altra equazione già vista, e cioè: $dQ/dt = a_1 Q + a_{11} Q^2$. La stessa equazione può descrivere anche certi casi di crescita organica e di relativo aumento di determinati organi e attività fisiologiche rispetto alle dimensioni del corpo. Ricordiamo infine, per concludere su questo punto, che anche il caso dell'associazione di due specie, di cui una si nutra dell'altra, è stato descritto dal Volterra con un sistema di equazioni che fa pure parte della descrizione generale di un sistema:

$$\begin{cases} \dfrac{dQ_1}{dt} = f_1(Q_1, Q_2)\,, \\ \dfrac{dQ_2}{dt} = f_2(Q_1, Q_2)\,. \end{cases}$$

Trascuriamo di esaminare lo sviluppo del sistema nel caso specifico considerato. Basta ricordare che ne derivano delle fluttuazioni regolari per cui l'aumento numerico della specie predatrice determina una riduzione della specie predata; onde, a causa della successiva mancanza di alimenti, si ha una riduzione anche della specie predatrice. Questo determina una ripresa della specie predata che fornisce nuovo alimento, e così via. L'esperienza ha confermato queste fluttuazioni.

Per concludere questo punto, ricordiamo ancora che le stesse equazioni sono state usate dallo Spiegelman ([21]) per la cinetica dei processi cellulari e per la teoria della competizione entro un organismo. Il Werner ([22]), infine, ha usato, come legge di base per l'azione farmacodinamica, un sistema di equazioni simili, ma più generali. Egli ha considerato infatti il sistema come con-

([20]) Per un prospetto generale di questi studi vedi: U. D'Ancona: *La lotta per l'esistenza* (Torino, 1942). Questo volume è stato preparato dall'autore proprio su suggerimento di von Bertalanffy.

([21]) S. Spiegelman: *Quart. Rev. Biol.*, **20**, 121 (1945).

([22]) G. Werner: *Sitz. Akad. Wiss. Wien Math.-Naturw. Kl.*, Abt. IIa. **156**, 457 (1947).

tinuo, usando perciò equazioni alle derivate parziali rispetto alle variabili x, y, z e t. Da tale legge di base si ricavano le varie leggi dell'azione della droga introducendo le condizioni speciali corrispondenti.

6. – Operatività del modello.

Non resta che esaminare l'ultima condizione cui era stato accennato: se cioè il modello può essere operativo, suscettibile cioè di trarne conseguenze verificabili. La risposta può esser data tenendo presente che il VON BERTALANFFY, dallo schema generale di cui si parla, ha ritenuto di trarre alcune conseguenze che hanno poi avuto una conferma sperimentale. Sulla operatività di questo modello del resto il VON BERTALANFFY non ha dubbi. « Ciò che è fondamentale per una vera biofisica, fisica dell'organismo vivente, è una teoria dei sistemi aperti e degli stati stazionari » ([23]), scrive.

Nel quadro della visione dell'organismo vivente come sistema aperto in stato stazionario, si collocano gli interessantissimi studi del VON BERTALANFFY tendenti a stabilire connessioni tra i due aspetti fondamentali degli organismi viventi: il metabolismo e l'accrescimento ([24]). Il metabolismo varia ovviamente con le dimensioni del corpo. Misurando il metabolismo in base al consumo di O_2 per chilogrammo di peso del corpo si era già notato che questo rapporto diminuiva coll'aumentare del peso del corpo e che il metabolismo aumentava circa in proporzione all'aumento della superficie, anzichè del peso. La superficie di un organismo (S) si può ricavare, come è noto, mediante la formula di Mech, dal peso (P) mediante la relazione approssimativa: $S = bP^{\frac{2}{3}}$ (*), dove b è un parametro legato alla particolare forma della specie considerata. Si hanno naturalmente anche formule rese più esatte mediante alcune correzioni, per esempio quella di Dubois valida nel caso dell'uomo, ma la (*) è sufficiente per il ragionamento che segue. Se il metabolismo (M) varia col variare della superficie, si avrà: $M = bP^{\frac{2}{3}}$. È facile vedere come questa formula si inquadri perfettamente nelle equazioni cinetiche dei sistemi. Consideriamo due elementi di un sistema, Q_1 e Q_2 nel caso più semplice in cui tutti i coefficienti sono uguali a zero tranne il primo, si avrà:

$$\begin{cases} \dfrac{dQ_1}{dt} = a_1 Q_1 , \\ \dfrac{dQ_2}{dt} = a_2 Q_2 , \end{cases} \qquad \text{da cui integrando:} \qquad \begin{cases} Q_1 = c_1 \exp[a_1 t] , \\ Q_2 = c_2 \exp[a_2 t] . \end{cases}$$

([23]) L. VON BERTALANFFY: *Scientia*, **11**, 361 (1954).

([24]) L. VON BERTALANFFY: *Quantitative Laws in Metabolism and Growth*. Ciclostilato senza indicazioni di luogo e di data, ma redatto nel 1956. Vedi anche, sempre di VON BERTALANFFY: *Nature*, 163 (1946) e *Amer. Nat.*, **85**, 111 (1951).

Risolvendo rispetto a t:

$$t = \frac{\log_e Q_1 - \log_e c_1}{a_1} = \frac{\log_e Q_2 - \log_e c_2}{a_2},$$

da cui

$$\log_e Q_1 = \frac{a_1}{a_2} \log_e Q_2 + \log_e c_1 - \frac{a_1}{a_2} \log_e c_2 ,$$

cioè

$$\log_e Q_1 = \log_e Q_2^{(a_1/a_2)} + \log_e \frac{c_1}{c_2^{(a_1/a_2)}},$$

e ponendo

$$A = \frac{a_1}{a_2} \qquad \text{e} \qquad B = \frac{c_1}{c_2^A},$$

e passando dai logaritmi ai numeri:

$$Q_1 = BQ_2^A ,$$

che corrisponde alla $M = bP^{\frac{2}{3}}$. In questo particolare caso $A = \frac{2}{3}$, ma non era evidentemente da escludere che questo valore non fosse universalmente valido, tanto più che la relazione metabolismo-superficie era in origine una relazione empirica, formulata in base a un numero relativamente ristretto di casi considerati. Il primo passo del VON BERTALANFFY fu di pensare ad una generalizzazione. Dal 1941 VON BERTALANFFY ed altri cercarono sperimentalmente come si comportasse la relazione considerata in molteplici gruppi animali e risultò che effettivamente l'esponente $\frac{2}{3}$ non valeva in tutti i casi. Valeva, salvo qualche correzione, per i mammiferi, i pesci, in genere per i vertebrati pecilotermi (per i quali, come è noto, non v'è termoregolazione) e per qualche invertebrato, ma non per gli altri animali. Si doveva perciò correggere la relazione: $M = bP^{\alpha}$. α assumeva valori tra $\frac{2}{3}$ e 1. Si potevano quindi delineare tre gruppi metabolici: il primo composto dagli animali con metabolismo in relazione alla superficie ($\alpha = \frac{2}{3}$), il secondo (larve di insetti, lumache, vermi) composto dagli animali in cui il metabolismo è invece in relazione direttamente al peso ($\alpha = 1$), il terzo (platelminti e alcune lumache) con valori del metabolismo intermedi a quelli dei due primi gruppi. L'interpretazione di questi tre tipi metabolici resta ancora aperta, comunque il VON BERTALANFFY utilizzò questi risultati per studiare l'accrescimento. La concezione dell'organismo come sistema aperto in stato stazionario dette l'avvio. Lo stato stazionario implica che il risultato globale di crescita o di equilibrio sia il risultato di processi dinamici continui, costruttivi e distruttivi. L'accrescimento deve perciò essere

il risultato della differenza tra i processi anabolici e quelli catabolici: $dP/dt = \eta P^m - kP^n$, dove η indica la costante dell'anabolismo, k la costante del catabolismo e m ed n indicano una qualche dipendenza col peso del corpo P. Questa concezione della crescita come contro-azione di anabolismo e di catabolismo, per cui, quando l'anabolismo non prevale più, si raggiunge lo stato stazionario, ha permesso, come si vedrà, di determinare i vari tipi di accrescimento in rapporto ai tipi metabolici sopra considerati. La relazione sopra scritta è evidentemente approssimata in notevole misura se si pensa alla semplificazione, che essa implica, di fenomeni che sono invece di una grande complessità. Il fatto che se ne traggano risultati soddisfacenti depone perciò a favore del modello di ragionamento.

Il VON BERTALANFFY osserva che i processi catabolici significano: biochimicamente, la degradazione dei materiali costruenti, particolarmente delle proteine, come è dimostrato dalle tecniche isotopiche; citologicamente, il rinnovo delle cellule. Il catabolismo si può pertanto assumere, in prima approssimazione, come direttamente proporzionale al peso. Perciò, anche considerando il fatto che l'equazione di cui sopra è poco sensibile a piccoli spostamenti di n dall'unità, si può porre $n = 1$.

La situazione per l'anabolismo è diversa. La sintesi dei componenti cellulari macromolecolari ha bisogno da un lato di costruire aminoacidi, zuccheri, fosfati, ecc. e dall'altro lato dell'energia fornita dai processi ossidativi. Del resto i risultati sperimentali indicano che, almeno negli animali superiori, vi è una regolare connessione tra respirazione, anabolismo e accrescimento. Perciò VON BERTALANFFY, con un'altra approssimazione, attribuisce all'esponente m i valori trovati per α nel caso dei tipi metabolici. Si avranno così, anche per l'accrescimento, tre casi,: $m = \frac{2}{3}$, $m = 1$, $\frac{2}{3} < m < 1$. Le predizioni dei tre tipi di accrescimento in funzione dei tipi metabolici sono state sperimentate e trovate corrette in gran numero di casi, spesso in modo insospettato. Nel primo caso la proporzione dell'accrescimento diminuisce progressivamente secondo una curva sigmoide, finchè si raggiunge eventualmente lo stato stazionario (mammiferi, con alcune restrizioni, pesci, alcuni invertebrati). Nel secondo caso (larve di insetti, ecc.) l'anabolismo ed il catabolismo procedono secondo lo stesso andamento e la proporzione dell'accrescimento tende ad aumentare indefinitamente (accrescimento esponenziale) senza raggiungere alcun stato stazionario. Il paradosso di un accrescimento indefinito è naturalmente solo apparente: l'accrescimento di una larva è infatti ad un certo punto tagliato dalla metamorfosi, quello di una lumaca di questo tipo dal ciclo stagionale, ecc. Nel terzo caso infine si hanno valori intermedi.

Dalle curve di accrescimento ottenute in relazione a questo modello possono ricavarsi i valori della costante catabolica k. Tali valori possono confrontarsi con quelli ottenuti sperimentalmente dalla distruzione delle proteine. La coincidenza tra i valori ottenuti con i due metodi, secondo VON BERTALANFFY, è

stata talora sorprendente, se si considera da un lato che il modello teorico di partenza è estremamente semplificato e dall'altro che l'errore nelle determinazioni fisiologiche della degradazione delle proteine è notevole. Ad esempio, nel 1938, il von BERTALANFFY calcolò il ritmo di degradazione delle proteine dalle curve di accrescimento dell'uomo, ottenendo il valore di 1.165 g/kg di peso del corpo al giorno. Nel 1949 SPRINSON e RITTENBERG fecero lo stesso calcolo partendo dai loro esperimenti con ^{15}N e trovarono il valore di $\frac{1}{3}$ g/kg di peso del corpo al giorno. L'accordo tra le due cifre non è molto, per la verità, ma tenendo conto della situazione è da considerare favorevolmente e quanto meno indicativo dell'ordine di grandezza del fenomeno. Non è qui il caso di discutere dettagliatamente questi problemi, nè la possibilità di apportare al modello correzioni, nè quelle già apportate, ma piuttosto di vedere altre conseguenze che il VON BERTALANFFY ha tratto dalla concezione dell'organismo come sistema aperto in stato stazionario e che sono state poi confermate.

Le considerazioni che precedono, ed altre, tutte legate alla teoria della crescita animale intesa come il risultato d'una azione antagonista di processi anabolici e catabolici ([25]) ha condotto il nostro autore ad interessanti conclusioni ([26]): 1) il catabolismo e la rigenerazione delle proteine hanno luogo a una velocità più grande di quanto non immagini la fisiologia classica ([27]); 2) si ha sintesi e risintesi di aminoacidi a partire da ammoniaca e da catene non azotate; 3) nel caso di una alimentazione povera in proteine, l'azoto proveniente dai costituenti dei corpi catabolizzati è trattenuto e riutilizzato per la rigenerazione proteica. Questo spiegherebbe l'azione di risparmio proteico giocata dagli idrati di carbonio, per cui una dieta ricca appunto in idrati di carbonio necessita di un minore apporto in proteine.

Scrive il VON BERTALANFFY: « Queste conclusioni sono state verificate da successivi lavori con isotopi, particolarmente con ^{15}N. La tecnica degli isotopi ha dimostrato il rapido rivolgimento delle proteine nel corpo animale, la rigenerazione delle proteine da un „metabolic pool" contenente azoto derivato dalle proteine degradate e l'utilizzazione dell'azoto ammoniacale del regime alimentare per la sintesi di proteine nelle diete povere di proteine » ([28]).

([25]) Il VON BERTALANFFY si sofferma su questa teoria particolarmente in: *Experientia*, **4**, 225 (1948) e nel saggio già citato in: *Nature* (Londra, 1949), p. 163.

([26]) L. VON BERTALANFFY: *Human Biology*, **10**, 181 (1938).

([27]) Il VON BERTALANFFY cita a conferma particolarmente: LEBLOND e STEVENS: *Anat. Rec.*, **100**, 357 (1948); STOREY e LEBLOND: *Ann. N.Y. Acad. Sci.*, **53**, 537 (1951); F. D. BERTALANFFY e LEBLOND: *Anat. Rec.*, **115**, 515 (1953); LEBLOND e WALKER: *Physiol. Rev.*, **36**, 225 (1956) e L. VON BERTALANFFY: *Biophysik des Fließgleichgewichts* (Braunschweig, 1953) per mostrare che anche il rinnovo delle cellule procede spesso ad un ritmo inaspettatamente alto.

([28]) L. VON BERTALANFFY: *Scientia*, **11**, 361 (1954).

596

Il VON BERTALANFFY cita a sostegno della sua tesi i lavori con ^{15}N di SPRINSON e RITTENBERG ([29]). Effettivamente questo continuo dinamismo negli organismi viventi, interessante anche sostanze in apparenza inerti come la dentina o l'osso, così lontano dall'equilibrio inteso in senso classico, ma pure tendente o mantenente uno stato stazionario, ci pare di non trascurabile peso in favore della teoria. Tanto più che questo ordine di idee ha permesso di dedurre altre conseguenze tendenti ad illustrare ulteriormente il lavoro d'insieme della cellula e dell'organismo. Partendo da queste premesse, infatti, KUHN ([30]) e VON BERTALANFFY ([31]) indicarono il lavoro minimo richiesto per produrre una molecola d'una sostanza polimera a concentrazione c diversa dalla concentrazione d'equilibrio c^* con la relazione: $W = RT \log c/c^*$. I calcoli di SCHULZ ([32]) esprimono approssimativamente il lavoro necessario per grammo di proteina con: $dw/dt = r(RT/Mw) \log(nw/K)$, ove r indica il rivolgimento di proteine per giorno, Mw il peso molecolare degli aminoacidi condensati in polimeri, nw il numero di molecole d'acqua per molecola di sostanza mono e polimera e K una costante d'equilibrio. Con opportuni valori ne deriva: $dw/dt = r \cdot 93$ cal/g proteina. Il VON BERTALANFFY attribuisce ad r valori medi ricavati dai lavori già citati di SPRINSON e RITTENBERG con l' ^{15}N, e ottiene che è necessario da uno a due per cento del metabolismo basale per mantenere le proteine dei corpi in stato stazionario. La sintesi proteica richiede dunque una ben modesta parte del metabolismo totale. Stime fatte da BORSOOK ([33]) e da LANG ([34]) con metodi diversi (esperienze con ^{15}N ma basate su ipotesi diverse e sulla duplicazione del materiale nucleare durante la mitosi) corrispondono ai dati ottenuti con il metodo sopra descritto sia pure a grandi linee. Commenta il VON BERTALANFFY: « Una conseguenza è che sarà futile cercare differenze nella respirazione cellulare tra organi con alta o bassa attività mitotica. Rispetto alla crescita maligna segue che un approccio dovrebbe esservi nella linea della sintesi proteica piuttosto che in quella del metabolismo dell'energia, come ha proposto RONDONI » ([35]).

Infine, prima di concludere questa serie di considerazioni sull'operatività del modello dell'organismo vivente come sistema aperto stazionario, ricorderemo un'ultima deduzione del VON BERTALANFFY non priva di interesse. Come è noto il DRIESCH considerò l'automoltiplicazione delle unità biologiche elementari (geni, cromosoni) come una seconda prova del vitalismo (la prima era

---

([29]) D. B. SPRINSON e D. RITTENBERG: *Journ. Biol. Chem.*, **180**, 707, 715 (1949).

([30]) W. KUHN: *Ergebn. Enzymforsch.*, **5**, 1 (1936).

([31]) L. VON BERTALANFFY: *Theoretische Biologie*, vol. **2**, già citato.

([32]) G. V. SCHULZ: *Naturwiss.*, **37**, 196 (1950).

([33]) H. BORSOOK: *Physiol. Rev.*, **30**, 206 (1950).

([34]) K. LANG: *Mikroskopische und chemische Organisation der Zelle*. 2. Colloquium d. deutschen Ges. f. physiol. Chemie (Berlin, 1952).

([35]) L. VON BERTALANFFY: *Scientia*, **11**, 361 (1954), già citato.

quella dell'« equifinalità » di cui si è già parlato). Ora il VON BERTALANFFY, partendo dalle sue idee, propose uno schema di autoduplicazione di tali unità biologiche elementari basato sull'intuizione che queste unità non fossero corpi stabili ed inerti, ma sistemi metabolizzati i quali, proprio mediante il metabolismo, si mantengono in uno stato stazionario crescendo per fissazione di gruppi molecolari appropriati su una matrice e subendo processi catabolici ([36]).

Lavori di LEBLOND e collaboratori con il radiofosforo ([37]) hanno mostrato che non si ha scambio di acido desossiribonucleico (DNA) nell'interfase, ma bensì un rimpiazzamento completo del DNA nello stadio premitotico, onde soltanto il DNA nuovo passa ai nuclei figli. Ciò sarebbe dimostrato dal fatto che nei tessuti con attività mitotica bassa, media, elevata (fegato, polmoni, mucosa intestinale) la quantità di DNA P formata è circa doppia di quella che corrisponderebbe al numero delle cellule neoformate. Ciò mostrerebbe appunto, secondo il VON BERTALANFFY, che il DNA è catabolizzato e rigenerato durante la mitosi. Ma non possiamo purtroppo entrare nel merito di questo problema peraltro quanto mai importante.

Questi cenni sono sufficienti, ci pare, per concludere sulla operatività del modello proposto. Ovviamente qui non interessava tanto entrare nel merito delle singole previsioni e delle relative conferme. Qui si voleva soltanto mostrare che dal modello si possono ricavare previsioni non solo qualitative, ma anche quantitative su singoli eventi, tali da poter essere sottoposte al confronto con l'esperienza che le verificherà o meno. Il modello generale proposto è pertanto operativo.

7. – Conclusione.

Il modello dell'organismo vivente come « sistema aperto stazionario », in base a quanto è stato detto nella Sezione 2 è dunque seriamente proponibile. Come si è detto la proponibilità soddisfa alcune condizioni fondamentali, ma non è di per sè un giudizio di rispondenza, di completezza, di validità. Essa pertanto non chiude, ma apre un discorso critico impegnato. Restano infatti da affrontare numerosi problemi. Innanzi tutto bisogna dare una risposta al quesito se il modello sia completo o no. Rispondere a questa domanda significa in fondo confrontare il modello con tutta la problematica biologica, oltre ai punti qui accennati. Si è fatto riferimento, per esempio, alla necessità di

([36]) L. VON BERTALANFFY: *Naturwiss.*, **32**, 46 (1944).

([37]) C. E. STEVENS, R. DAOUST e C. P. LEBLOND: *Journ. Biol. Chem.*, **202**, 177 (1953); C. E. STEVENS, R. DAOUST e C. P. LEBLOND: *Can. Journ. Med. Sci.*, **31**, 263 (1953); R. DAOUST, F. D. BERTALANFFY e C. P. LEBLOND: *Journ. Biol. Chem.*, **207**, 405 (1954).

508

esaminare la corrispondenza tra il modello proposto e la genetica, e, soprattutto, i problemi legati alla modalità con cui il gene agisce nel determinare i caratteri che controlla, ma si potrebbero aggiungere altri punti su cui certamente è fondamentale una verifica dello schema: dai problemi della contrazione del muscolo ai modelli matematici della cellula di Rashevsky; dai più recenti studi di embriologia a tutta la vasta problematica della biochimica, ecc.

E non si è parlato di tutti gli aspetti più strettamente microfisici, di così grande interesse per la biologia anche se ci si volesse limitare, nella zona che sta tra il livello molecolare e quello atomico, ai loro aspetti teorici più generali quali quelli suggeriti panoramicamente dal MONDELLA nel suo bel lavoro recente ([38]).

In ultima analisi, del resto, solo quando i Q, gli X e gli J delle equazioni esaminate potranno acquistare dei significati e dei valori ben precisi la teoria avrà assolto il suo compito.

Qui intanto volevamo solo segnalare l'interesse di questo ordine di idee e mostrare come oggi siano possibili in biologia, in linea di massima, modelli teorici tali da promuovere la ricerca e da rendere utilizzabile il formalismo matematico.

([38]) F. MONDELLA: *Il Pensiero*, **3** (1957).

SUPPLEMENTO AL VOLUME XII, SERIE X
DEL NUOVO CIMENTO

N. 1, 1959
2° Trimestre

EMENDAMENTI ALLA NOTA

On the Origin of the Charge Spectrum of the Primary Cosmic Radiation.

pubblicata a pagg. 548-555 del *Supplemento*
al Vol. VIII, Serie X, del *Nuovo Cimento*, 1958

S. F. SINGER
University of Maryland - College Park, Maryland

(ricevuto il 10 Aprile 1959)

1. – Sostituire la Tabella I a pag. 550 con la seguente.

TABLE I. – *Data on cosmic abundances, including the primary cosmic radiation (beyond $Z=10$, only the most abundant species are listed).*

Z		a	b	c	d
1	H	$4 \cdot 10^{10}$	12.00	620	180
2	He	$3 \cdot 10^{9}$	11.06	90	20
3	Li	100	0.86		
4	Be	20	2.43	3	~ 0.4
5	B	24	[5.0]		
6	C	$3.5 \cdot 10^{6}$	8.56		
7	N	$6.6 \cdot 10^{6}$	7.98	6	0.8
8	O	$2.15 \cdot 10^{7}$	9.00		
9	F	1600			
10	Ne	$8.6 \cdot 10^{6}$	8.81		
14	Si	$1 \cdot 10^{6}$	7.60		
18	A	$1.5 \cdot 10^{5}$	7.7	2.5	0.3
20	Ca	$4.9 \cdot 10^{4}$	6.38		
26	Fe	$6 \cdot 10^{5}$	6.76		

(*a*) H. E. SUESS and H. C. UREY, *Rev. Mod. Phys.* **28**, 53 (1956). (These values are referred to silicon $= 1 \cdot 10^{6}$).

(*b*) log N for the sun, except that for He and Ne we give their ratios to H for stars (after ALLER).

(*c*) Vertical cosmic ray flux (in particles per m² ster s (peters) at $\lambda = 41°$ [8].

(*d*) Vertical cosmic ray flux (in peters) at $\lambda = 0°$ [8].

2· – Sostituire le References di pag. 555 con le seguenti.

REFERENCES

[1] B. Peters in J. G. Wilson's: *Progress in Cosmic Ray Physics*, vol. I (Amsterdam, 1952).
[2] P. Morrison, S. Olbert and B. Rossi: *Phys. Rev.*, **94**, 440 (1954).
[3] B. Rossi: *Suppl. Nuovo Cimento*, **2**, 275 (1955).
[4] A. D. Dainton, P. H. Fowler and D. W. Kent: *Phil. Mag.*, **43**, 317 (1952).
[5] S. F. Singer: *Proc. Duke Univ. Cosmic Ray Conf.*, Durham, No. Carol. (Nov. 1953), ed. by L. Nordheim (1954).
[6] S. F. Singer: *Phys. Rev.*, A **98**, 1163 (1955).
[7] S. Hayakawa: *Progr. Theor. Phys.*, **15**, 111 (1956).
[8] S. F. Singer: in *Progress in Cosmic Ray Physics*, vol. **4**, (Amsterdam, 1957).
[9] P. H. Fowler, R. R. Hillier and C. J. Waddington: *Phil. Mag.*, (1957).
[10] J. H. Noon and M. F. Kaplon: *Phys. Rev.*, **97**, 769 (1955).
[11] E. L. Fireman: *Phys. Rev.*, **97**, 1303 (1955); L. A. Currie, W. F. Libby and R. L. Wolfgang: *Phys. Rev.*, **101**, 1557 (1956).
[12] F. Begemann, J. Geiss and D. C. Hess: *Phys. Rev.*, **107**, 540 (1957).
[12] F. Begemann, J. Geiss and D. C. Hess: *Phys. Rev.*, **107**, 540 (1957); S. F. Singer: *Phys. Rev.* (in press).
[13] C. Dahanayake *et al.* (Bristol): private communication.

PROPRIETÀ LETTERARIA RISERVATA

Direttore responsabile: G. POLVANI

Tipografia Compositori - Bologna

Questo fascicolo del *Supplemento* è stato licenziato dai torchi il 21-VII-1959